



TESIS - TE142599

**REKONSTRUKSI OBJEK 3D PADA VIDEO STEREO  
DENGAN *FOCAL LENGTH* DINAMIS  
BERBASIS FUNDAMENTAL MATRIX**

KURNIA PRIMA PUTRA  
2215205201/07111550052001

DOSEN PEMBIMBING  
Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.  
Dr. Eko Mulyanto Yuniarno, ST., MT.

PROGRAM MAGISTER  
BIDANG KEAHLIAN JARINGAN CERDAS MULTIMEDIA  
DEPARTEMEN TEKNIK ELEKTRO  
FAKULTAS TEKNOLOGI ELEKTRO  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2018





TESIS - TE142599

**REKONSTRUKSI OBJEK 3D PADA VIDEO STEREO  
DENGAN *FOCAL LENGTH* DINAMIS  
BERBASIS FUNDAMENTAL MATRIX**

KURNIA PRIMA PUTRA  
2215205201/07111550052001

DOSEN PEMBIMBING  
Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.  
Dr. Eko Mulyanto Yuniarno, ST., MT.

PROGRAM MAGISTER  
BIDANG KEAHLIAN JARINGAN CERDAS MULTIMEDIA  
DEPARTEMEN TEKNIK ELEKTRO  
FAKULTAS TEKNOLOGI ELEKTRO  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2018





## LEMBAR PENGESAHAN

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar  
Magister Teknik (M.T)  
di  
Institut Teknologi Sepuluh Nopember

oleh:

Kurnia Prima Putra  
NRP. 2215205201 / 07111550052001

Tanggal Ujian : 26 Juni 2018  
Periode Wisuda : September 2018

Disetujui oleh:

1. Prof. Dr. Ir. Mauridhi Hery P., M.Eng. (Pembimbing I)  
NIP : 195809161986011001

2. Dr. Eko Mulyanto Yuniarno, ST., MT. (Pembimbing II)  
NIP : 196806011995121009

3. Dr. I Ketut Eddy Purnama, ST., MT. (Penguji)  
NIP : 196907301995121001

4. Dr. Surya Sumpeno, ST., M.Sc. (Penguji)  
NIP : 196906131997021003

Dekan Fakultas Teknologi Elektro

Dr. Tri Arief Sardjono, ST., MT.  
NIP. 197002121995121001



## PERNYATAAN KEASLIAN TESIS

Dengan ini saya menyatakan bahwa isi keseluruhan Tesis saya dengan judul **REKONSTRUKSI OBJEK 3D PADA VIDEO STEREO DENGAN *FOCAL LENGTH* DINAMIS BERBASIS FUNDAMENTAL MATRIX** adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diijinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri.

Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, 26 Juni 2018

Kurnia Prima Putra

2215205201 / 07111550052001

*Halaman ini sengaja dikosongkan*

# REKONSTRUKSI OBJEK 3D PADA VIDEO STEREO DENGAN *FOCAL LENGTH* DINAMIS BERBASIS FUNDAMENTAL MATRIX

Nama Mahasiswa : Kurnia Prima Putra  
NRP : 2215205201 / 07111550052001  
Pembimbing :1. Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.  
2. Dr. Eko Mulyanto Yuniarno, ST., MT.

## ABSTRAK

Proses rekonstruksi tiga dimensi (3D) citra dapat didefinisikan sebagai proses mengolah citra-citra dua dimensi menjadi bentuk 3D yang lebih konstruktif dan membantu pengamat dalam menganalisa informasi-informasi yang terkandung di dalam citra. Rekonstruksi 3D menggunakan citra yang berasal dari kamera digital memerlukan pengambilan beberapa citra suatu objek atau lingkungan dari titik dan/atau sudut yang berbeda. Salah satu jenis citra dari sudut pandang yang berbeda adalah citra stereo. Pada pengambilan citra jenis ini, transformasi antar kamera yang terjadi berupa translasi murni. Karena transformasi berupa translasi murni, maka estimasi parameter ekstrinsik suatu kamera dapat ditentukan dengan melakukan translasi pada parameter ekstrinsik kamera lain.

Pada citra stereo, pencarian dan penentuan korespondensi piksel-piksel pada citra kiri ke citra kanan merupakan hal yang penting untuk melakukan rekonstruksi 3D objek atau lingkungan. Permasalahan pada proses ini adalah ketelitian penentuan korespondensi pada level piksel karena *pixel locking*. Hal ini mengakibatkan detail dari rekonstruksi menurun. Salah satu metode yang digunakan untuk meningkatkan ketelitian adalah dengan mengambil citra dengan tingkat *zoom* yang berbeda-beda. Perbedaan *zoom* mengakibatkan perubahan pada *focal length* sehingga korespondensi titik-titik antar citra juga ikut berubah. Selain itu, perubahan *zoom* juga mengakibatkan pergeseran *principal point* pada masing-masing citra. Pada beberapa penelitian sebelumnya, perhitungan *zoom factor* dilakukan dengan melihat piksel-piksel fitur yang saling berkorespondensi pada citra-citra dari kamera yang sama dengan *zoom berbeda*. Cara ini sangat sensitif dengan korespondensi koordinat titik pada tingkat subpiksel sehingga area simpangan saat melakukan estimasi titik pusat ekspansi meluas. Pada penelitian ini, rekonstruksi objek 3D dilakukan dengan memanfaatkan fundamental matriks untuk mendapatkan *focal length* dan memperkecil area simpangan estimasi titik pusat ekspansi.

Keywords : Rekonstruksi 3D, citra stereo, fundamental matriks, *focal length* dinamis

*Halaman ini sengaja dikosongkan*

# 3D OBJECT RECONSTRUCTION ON STEREO CAMERA WITH DYNAMIC FOCAL LENGTH BASED ON FUNDAMENTAL MATRIX

By : Kurnia Prima Putra  
Student ID Number : 2215205201 / 07111550052001  
Advisors :1. Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.  
2. Dr. Eko Mulyanto Yuniarno, ST., MT.

## ABSTRACT

Image three dimensional reconstruction processs could be defined as processing two dimensional image to form three dimensional object which more constructive and helps observers to analyze informations from images. 3D reconstruction used image from digital camera requires to take several images of an object or environment from different position and/or viewing angles. An example of images taken from different position is stereo images. Transformation between camera when taking this images is pure translation. Since the transformation is pure translation, extrinsic parameters of a camera could be obtained by translatinng the extrinsic parameters of other camera.

On stereo images, finding and determining pixel correspondences of leftside images to rightside images is an important step to accomplish a 3D object or environment. The problem of this process is the precision of correspondences are at pixwl level due to pixel lock. This caused the detail of reconstruction lowered. One of the method used to increase the reconstruction detail is to capture on image with different zoom levels. Changes on zoom caused the changes on focal length thus points correspondeces also changes. Furthermore, zoom changes also caused shifts on principal point on each images. On previous researches, the calculation of *zoom factor* done by observing corresponding featured pixels of images from same camera with different zoom. This method quite sensitive to coordinate correspondence at subpixel level this the error area when estimating the center of expansion mmight get wider. In this research, we try to reconsruct 3D object by exploiting fundamental matrix to obtain focal length and to reduce estimated error area of center of expansion.

Keywords : 3D Reconstruction, stereo images, fundamental matrix, dynamic focal length

*Halaman ini sengaja dikosongkan*



## KATA PENGANTAR

Dengan nama Allah Yang Maha Pengasih dan Penyayang. Puji syukur kehadirat Allah SWT atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tesis yang berjudul Rekonstruksi objek 3D pada Video Stereo dengan *focal length* Dinamis berbasis Fundamental matriks. terselesaikannya tesis ini tentu berkat dukungan berupa semangat dan doa dari berbagai pihak yang telah terlibat baik secara langsung maupun tidak langsung. Oleh karena itu penulis mengucapkan terima kasih yang tidak terhingga kepada :

1. Kedua orang tua Bpk. Pudji Mulyo Winarno (alm) dan Ibu Ety Tedjo Dwicahyowati yang terus memberikan motivasi dan menyemangati penulis untuk menyelesaikan tesis.
2. Terima kasih yang tak terhingga kepada Dosen Pembimbing Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng. dan Dr. Eko Mulyanto Yuniarno, ST., MT., yang telah memberikan bimbingan dan masukan dalam pengerjaan tesis dari awal hingga akhir.
3. Terima kasih kepada seluruh dosen Jaringan Cerdas Multimedia atas kesediaan membagi ilmu dan pengalamannya kepada penulis.
4. Terima kasih untuk sahabat-sahabat penulis dan semua rekan-rekan di JCM yang telah menemani dan berjuang bersama hingga menyelesaikan studi di jenjang perguruan tinggi ini yang tidak bisa penulis sebutkan satu per-satu.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna, oleh karena itu kritik dan saran yang bersifat membangun sangat diharapkan untuk menyempurnakan tesis ini. Penulis berharap skripsi ini dapat bermanfaat bagi perkembangan teknologi rekonstruksi 3D dan *3D sensing*.

Surabaya, Mei 2018

Penulis

*Halaman ini sengaja dikosongkan*

# DAFTAR ISI

LEMBAR PENGESAHAN.....	iii
PERNYATAAN KEASLIAN TESIS.....	v
ABSTRAK .....	vii
ABSTRACT .....	ix
KATA PENGANTAR.....	xi
DAFTAR GAMBAR.....	xv
DAFTAR TABEL .....	xxiii
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Masalah.....	4
1.5 Kontribusi .....	4
BAB 2 KAJIAN PUSTAKA.....	5
2.1 Penelitian Terkait .....	5
2.1.1 Kajian Penelitian Terkait.....	5
2.2 Teori Dasar .....	6
2.2.1 Rekonstruksi Tiga Dimensi Citra .....	7
2.2.2 Citra Stereo .....	9
2.2.3 <i>Zooming</i> .....	12
BAB 3 METODOLOGI PENELITIAN .....	15
3.1 Kalibrasi.....	15
3.2 Akuisisi data .....	16
3.3 Estimasi <i>focal length</i> otomatis.....	17
3.3.1 Estimasi <i>focal length</i> berbasis jarak titik.....	18
3.3.2 Estimasi <i>focal length</i> menggunakan fundamental matriks ..	20
3.4 Normalisasi koordinat titik .....	23
3.4.1 Estimasi <i>center of expansion</i> berbasis garis ekspansi.....	25
3.4.2 Estimasi <i>center of expansion</i> menggunakan fundamental matriks.....	26

3.5	Rekonstruksi Tiga Dimensi .....	31
3.6	Evaluasi hasil.....	32
BAB 4	HASIL DAN PEMBAHASAN.....	35
4.1	Kalibrasi.....	35
4.1.1	Spesifikasi kamera virtual.....	35
4.1.2	Spesifikasi kamera fisik .....	36
4.2	Akuisi data.....	37
4.2.1	Akuisisi data sintetis.....	37
4.2.2	Akuisisi data citra.....	39
4.3	Estimasi <i>focal length</i> otomatis.....	40
4.3.1	Estimasi <i>focal length</i> berbasis jarak titik .....	41
4.3.2	Estimasi <i>focal length</i> menggunakan fundamental matriks ..	53
4.4	Normalisasi koordinat titik .....	65
4.4.1	Normalisasi berbasis garis ekspansi .....	65
4.4.2	Normalisasi menggunakan fundamental matriks.....	72
4.5	Rekonstruksi Tiga Dimensi .....	80
4.5.1	Rekonstruksi tanpa fundamental matriks .....	80
4.5.2	Rekonstruksi memanfaatkan fundamental matriks.....	86
4.6	Evaluasi Hasil .....	92
4.6.1	Evaluasi pada data sintetis .....	93
4.6.2	Evaluasi pada data citra.....	104
BAB 5	KESIMPULAN .....	107
DAFTAR PUSTAKA	.....	109
BIOGRAFI PENULIS	.....	111

## DAFTAR GAMBAR

2.1	Model <i>pinhole camera</i> . . . . .	8
2.2	Triangulasi (Hartley & Zisserman 2002) . . . . .	9
2.3	Geometri Epipolar (Bhatti & Navahandi 2011). . . . .	12
2.4	<i>Zooming</i> pada model <i>pinhole camera</i> . . . . .	13
3.1	Metodologi penelitian . . . . .	15
3.2	Pola papan catur yang digunakan . . . . .	16
3.3	Perubahan posisi titik-titik fitur pada perubahan <i>zoom</i> . . . . .	18
3.4	Tiga pasangan citra yang digunakan dalam estimasi <i>focal length</i> menggunakan fundamental matriks . . . . .	20
3.5	Perubahan garis <i>epipolar</i> pada <i>focal length</i> berbeda . . . . .	21
3.6	Proses normalisasi menggunakan <i>inverse</i> dari <i>zooming</i> . . . . .	23
3.7	Garis-garis ekspansi berpotongan pada titik <i>center of expansion</i> . . . . .	25
3.8	Perpotongan garis-garis ekspansi . . . . .	27
3.9	Perpotongan garis-garis ekspansi dengan garis $l_c$ . . . . .	31
4.1	Model virtual tiga dimensi . . . . .	37
4.2	Hasil proyeksi model tiga dimensi virtual. . . . .	38
4.3	Data <i>scene</i> pertama, objek dengan pola papan catur. . . . .	39
4.4	Data <i>scene</i> kedua, objek balok pada lingkungan dalam ruangan. . . . .	40
4.5	Data <i>scene</i> ketiga, objek botol pada lingkungan dalam ruangan. . . . .	40
4.6	Data <i>scene</i> keempat, lingkungan luar ruangan. . . . .	41
4.7	Hasil $s_n$ berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 1. . . . .	42
4.8	Hasil estimasi <i>focal length</i> berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 1. . . . .	42
4.9	Hasil $s_n$ berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 2. . . . .	43
4.10	Hasil estimasi <i>focal length</i> berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 2. . . . .	44

4.11 Hasil $s_n$ berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3. . . . .	45
4.12 Hasil estimasi <i>focal length</i> berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3. . . . .	45
4.13 Hasil $s_n$ berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis. . . . .	46
4.14 Hasil estimasi <i>focal length</i> berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis. . . . .	47
4.15 Hasil estimasi $s_n$ berbasis jarak titik pada <i>scene</i> pertama. . . . .	48
4.16 Hasil estimasi <i>focal length</i> berbasis jarak titik pada <i>scene</i> pertama. . . . .	48
4.17 Hasil estimasi $s_n$ berbasis jarak titik pada <i>scene</i> kedua. . . . .	49
4.18 Hasil estimasi <i>focal length</i> berbasis jarak titik pada <i>scene</i> kedua. . . . .	50
4.19 Hasil estimasi $s_n$ berbasis jarak titik pada <i>scene</i> ketiga. . . . .	50
4.20 Hasil estimasi <i>focal length</i> berbasis jarak titik pada <i>scene</i> ketiga. . . . .	51
4.21 Hasil estimasi $s_n$ berbasis jarak titik pada <i>scene</i> keempat. . . . .	52
4.22 Hasil estimasi <i>focal length</i> berbasis jarak titik pada <i>scene</i> keempat. . . . .	52
4.23 Hasil $s_n$ menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1. . . . .	54
4.24 Hasil estimasi <i>focal length</i> menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1. . . . .	54
4.25 Hasil $s_n$ menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2. . . . .	55
4.26 Hasil estimasi <i>focal length</i> menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2. . . . .	56
4.27 Hasil $s_n$ menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3. . . . .	57

4.28	Hasil estimasi <i>focal length</i> menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3. . . . .	57
4.29	Hasil $s_n$ menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis. . . . .	58
4.30	Hasil estimasi <i>focal length</i> menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis. . . . .	59
4.31	Hasil estimasi $s_n$ menggunakan fundamental matriks pada <i>scene</i> pertama. . . . .	60
4.32	Hasil estimasi <i>focal length</i> menggunakan fundamental matriks pada <i>scene</i> pertama. . . . .	61
4.33	Hasil estimasi $s_n$ menggunakan fundamental matriks pada <i>scene</i> kedua. . . . .	61
4.34	Hasil estimasi <i>focal length</i> menggunakan fundamental matriks pada <i>scene</i> kedua. . . . .	62
4.35	Hasil estimasi $s_n$ menggunakan fundamental matriks pada <i>scene</i> ketiga. . . . .	63
4.36	Hasil estimasi <i>focal length</i> menggunakan fundamental matriks pada <i>scene</i> ketiga. . . . .	63
4.37	Hasil estimasi $s_n$ menggunakan fundamental matriks pada <i>scene</i> keempat. . . . .	64
4.38	Hasil estimasi <i>focal length</i> menggunakan fundamental matriks pada <i>scene</i> keempat. . . . .	65
4.39	Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 1. . . . .	66
4.40	Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 2. . . . .	67
4.41	Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3. . . . .	68

4.42 Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis.	69
4.43 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> pertama. . . . .	70
4.44 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> kedua. . . . .	71
4.45 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> ketiga. . . . .	71
4.46 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> keempat. . . . .	72
4.47 Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 1. . . . .	73
4.48 Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 2. . . . .	74
4.49 Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3. . . . .	75
4.50 Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis.	76
4.51 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> pertama. . . . .	77
4.52 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> kedua. . . . .	78
4.53 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> ketiga. . . . .	79
4.54 Hasil normalisasi berbasis garis ekspansi dari data citra <i>scene</i> keempat. . . . .	79
4.55 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1.	81
4.56 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2.	81



4.57 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3.	82
4.58 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis. . . . .	83
4.59 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra <i>scene</i> pertama. . . . .	84
4.60 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra <i>scene</i> kedua. . . . .	84
4.61 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra <i>scene</i> ketiga. . . . .	85
4.62 Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra <i>scene</i> keempat. . . . .	86
4.63 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1.	87
4.64 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2.	87
4.65 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3.	88
4.66 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan <i>image center</i> dinamis. . . . .	89
4.67 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra <i>scene</i> pertama. . . . .	90
4.68 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra <i>scene</i> kedua. . . . .	91
4.69 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra <i>scene</i> ketiga. . . . .	91
4.70 Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra <i>scene</i> keempat. . . . .	93

4.71	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi penuh. . . . .	93
4.72	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.01 piksel. . . . .	94
4.73	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.1 piksel. . . . .	94
4.74	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 1 piksel. . . . .	94
4.75	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi penuh. . . . .	95
4.76	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.01 piksel. . . . .	95
4.77	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.1 piksel. . . . .	95
4.78	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 1 piksel. . . . .	95
4.79	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi penuh. . . . .	96
4.80	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.01 piksel. . . . .	96
4.81	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.1 piksel. . . . .	97
4.82	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 1 piksel. . . . .	97
4.83	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi penuh. . . . .	97
4.84	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.01 piksel. . . . .	98
4.85	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.1 piksel. . . . .	98

4.86	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 1 piksel. . . . .	98
4.87	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi penuh. . . . .	99
4.88	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.01 piksel. . . . .	99
4.89	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.1 piksel. . . . .	99
4.90	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 1 piksel. . . . .	99
4.91	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi penuh. . . . .	100
4.92	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.01 piksel. . . . .	100
4.93	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.1 piksel. . . . .	100
4.94	Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 1 piksel. . . . .	101
4.95	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi penuh. . . . .	101
4.96	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi 0.01 piksel. . . . .	101
4.98	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi 1 piksel. . . . .	102
4.97	Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi 0.1 piksel. . . . .	102

4.99 Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi penuh. . . . .	103
4.100 Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi 0.01 piksel. . . . .	103
4.101 Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi 0.1 piksel. . . . .	103
4.102 Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan <i>image center</i> dinamis pada simulasi akurasi 1 piksel. . . . .	103
4.103 Perbandingan titik-titik hasil rekonstruksi pada scene pertama.	104
4.104 Perbandingan titik-titik hasil rekonstruksi pada scene kedua. .	105
4.105 Perbandingan titik-titik hasil rekonstruksi pada scene ketiga. .	105
4.106 Perbandingan titik-titik hasil rekonstruksi pada scene keempat.	106

## DAFTAR TABEL

4.1	Parameter intrinsik pada kamera stereo virtual pertama . . . .	35
4.2	Parameter intrinsik pada kamera stereo virtual kedua . . . . .	36
4.3	Parameter intrinsik pada kamera stereo virtual ketiga . . . . .	36
4.4	Parameter intrinsik pada <i>zoom</i> minimal . . . . .	36
4.5	Spesifikasi simulasi proses <i>zooming</i> pada data sintetis . . . . .	38

*Halaman ini sengaja dikosongkan*

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Proses rekonstruksi tiga dimensi (3D) citra dapat didefinisikan sebagai proses mengolah citra-citra dua dimensi menjadi bentuk 3D yang lebih konstruktif dan membantu pengamat dalam menganalisa informasi-informasi yang terkandung di dalam citra. Rekonstruksi 3D citra sangat penting pada bidang kedokteran, game dan robotika. Secara umum, teknik rekonstruksi terbagi menjadi dua jenis antara lain metode aktif dan pasif. Pada metode aktif, sistem mengukur objek secara aktif baik secara mekanis maupun radiometri. Hal ini dilakukan dengan mengirimkan sinyal pengukur radio (elektromagnetik) atau mekanis (audio) kemudian menganalisa sinyal yang dipantulkan oleh objek. Teknik yang tersebut memerlukan perangkat khusus yang cukup mahal sehingga mempersulit implementasi pada hal-hal baru. Pada metode pasif, sistem menganalisa sinyal-sinyal yang dipantulkan objek tanpa mengirimkan sinyal pengukur. Sistem ini dapat melibatkan penggunaan sensor-sensor yang sensitif terhadap cahaya inframerah atau cahaya tampak seperti pada kamera digital pada umumnya. Metode pasif relatif lebih aplikatif karena perangkat yang digunakan relatif mudah didapat dan ditemukan. Hal ini mendukung penerapan-penerapan baru pada bidang komputer grafis, virtual reality dan komunikasi yang kebutuhannya dalam rekonstruksi 3D meningkat beberapa tahun terakhir (Khilar et al. 2013).

Rekonstruksi 3D menggunakan citra yang berasal dari kamera digital memerlukan pengambilan beberapa citra suatu objek atau lingkungan dari titik dan/atau sudut yang berbeda. Salah satu jenis citra dari sudut pandang yang berbeda adalah citra stereo. Pada kamera stereo, transformasi antar kamera berupa translasi murni. Karena transformasi berupa translasi murni, maka estimasi parameter ekstrinsik suatu kamera dapat ditentukan dengan melakukan translasi pada parameter ekstrinsik kamera lain. Selain itu, estimasi fundamental matriks yang digunakan untuk representasi epipolar pada

citra stereo lebih sederhana. Estimasi fundamental matriks pada pergerakan tak tentu memiliki 7 *degree of freedom* sedangkan pada citra stereo hanya memiliki 2 *degree of freedom* (Hartley & Zisserman 2002). Garis-garis epipolar yang dibangun dari fundamental matriks citra stereo berupa garis-garis horizontal paralel (Nalpantidis et al. 2009). Karena karakteristik citra stereo ini, dua dekade terakhir stereo vision dianggap metode yang layak serta low-cost untuk mendapatkan informasi kedalaman (*depth*) pada beragam lingkungan dan aplikasi (Vancea et al. 2016).

Untuk mendapatkan hasil rekonstruksi objek yang lebih detail diperlukan citra objek dengan resolusi yang lebih tinggi. Perubahan resolusi citra tangkapan kamera dapat dilakukan melalui *zooming* lensa (Li & Lavest 1996). Proses *zooming* memberikan efek perbesaran atau pengecilan pada citra yang terekam kamera (Hartley & Zisserman 2002). Dalam model *pinhole camera*, proses *zooming* mengakibatkan perubahan pada parameter intrinsik kamera termasuk *focal length* (Willson 1994). Estimasi parameter intrinsik lebih mudah dilakukan dengan menggunakan objek khusus seperti pola papan catur (Zollner & Sablatnig 2004). Namun selama proses *zooming*, mengkalibrasi kamera secara menggunakan objek khusus seperti pola papan catur setiap frame untuk mendapatkan *focal length* merupakan hal yang kurang praktis untuk diterapkan. Untuk mengatasi hal ini proses *self-calibration* perlu dilakukan.

*Self-calibration* adalah proses mendapatkan parameter-parameter intrinsik kamera dengan memanfaatkan kumpulan citra tanpa objek khusus seperti pola papan catur. Proses ini sangat penting untuk mengatasi permasalahan-permasalahan kalibrasi ulang pada kamera yang mengalami perubahan *zoom* sangat sering (Elamsy et al. 2014). Penelitian-penelitian sebelumnya *self-calibration* untuk mendapatkan parameter intrinsik dilakukan pada kamera tunggal (Elamsy et al. 2014, M. Pollefeys & Gool 1998). Dengan demikian, *self-calibration* pada kamera stereo dilakukan pada masing-masing kamera secara terpisah. Hal ini tentunya memerlukan korespondensi titik-titik antar frame pada masing-masing kamera sehingga pada setiap dua frame, pencarian



korespondensi untuk proses rekonstruksi pada 4 pasang citra. Pada penelitian ini, hasil rekonstruksi *zoom* dinamis dari citra stereo dengan estimasi *focal length* pada pasangan citra stereo dengan memanfaatkan fundamental matriks dibandingkan terhadap hasil rekonstruksi dengan estimasi *focal length* yang dilakukan terpisah antar kamera.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang masalah, penelitian ini merumuskan beberapa poin-poin permasalahan antara lain :

1. Pendekatan *zoom* dinamis digunakan untuk meningkatkan detail dari objek
2. Perubahan *zoom* merubah merubah *focal length*
3. Perubahan *focal length* merubah relasi antara titik-titik di dunia nyata dengan titik-titik pada citra
4. Estimasi nilai *focal length* untuk kamera stereo dilakukan secara terpisah sehingga memerlukan proses pencarian korespondensi tambahan antar citra yang diambil pada kamera yang sama.

## 1.3. Tujuan Penelitian

Dari permasalahan yang telah dirumuskan, penelitian ini dilakukan dengan tujuan antara lain :

1. Melakukan rekonstruksi objek tiga dimensi dengan memanfaatkan informasi yang didapat dari fundamental matriks
2. Mengurangi proses korespondensi yang diperlukan dalam proses rekonstruksi dengan melakukan estimasi *focal length* memanfaatkan fundamental matriks

3. Membandingkan hasil rekonstruksi dengan estimasi *focal length* berbasis fundamental matriks terhadap estimasi *focal length* yang dilakukan terpisah antar kamera

#### **1.4. Batasan Masalah**

Pada penelitian ini peneliti membatasi ruang lingkup permasalahannya yaitu sebagai berikut :

1. Pengamatan pada citra stereo bergerak
2. Variasi yang terjadi hanya pada *zoom*
3. Lingkungan atau objek yang diamati bersifat statis

#### **1.5. Kontribusi**

Dengan memanfaatkan penelitian ini, penelitian berikutnya diharapkan meningkatkan performa rekonstruksi tiga dimensi maupun *3d sensing* yang menggunakan citra stereo tanpa melalui proses kalibrasi ketika mengalami proses *zooming*.

## BAB 2

### KAJIAN PUSTAKA

Dalam penelitian ini, beberapa kajian pustaka akan dilakukan. Berikut adalah kajian-kajian yang dilakukan dalam menjalankan penelitian ini.

#### 2.1. Penelitian Terkait

Pada penelitian ini ada beberapa penelitian terkait yang berusaha meningkatkan performa *stereo algorithm* :

1. Cristian-Cosmin Vancea(2016), Improving stereo reconstruction by sub-pixel correction using histogram matching, IEEE Intelligent Vehicles Symposium, Proceedings, Gothenburg, Sweden.
2. T. Elamsy, A. Habed, and B. Boufama, Self-calibration of stationary non-rotating zooming cameras, Image Vis. Comput., vol. 32, no. 3, pp. 212226, 2014.
3. R.G. Willson, Modeling and Calibration of Automated Zoom Lenses, (Ph.D. thesis) CarnegieMellon University, Pittsburgh, PA, 1994

Dari penelitian tersebut menjadi dasar studi literatur untuk memahami teknik-teknik serta *state of the art* dari *stereo algorithm*.

##### 2.1.1. Kajian Penelitian Terkait

Dua dekade terakhir stereo vision dianggap metode yang layak serta low-cost untuk mendapatkan informasi kedalaman (depth) pada beragam lingkungan dan aplikasi (Vancea et al. 2016). Pada pengambilan citra jenis ini, transformasi antar kamera yang terjadi berupa translasi murni. Karena transformasi berupa translasi murni, maka estimasi parameter ekstrinsik suatu kamera dapat ditentukan dengan melakukan translasi pada parameter ekstrinsik kamera lain. Selain itu, estimasi fundamental matriks yang digunakan untuk representasi epipolar pada citra stereo lebih sederhana. Estimasi fundamental matriks pada pergerakan tak tentu memiliki 7 degree of freedom sedangkan

pada citra stereo hanya memiliki 2 degree of freedom (Hartley & Zisserman 2002).

Untuk mendapatkan hasil rekonstruksi objek yang lebih detail diperlukan citra objek dengan resolusi yang lebih tinggi. Perubahan resolusi citra tangkapan kamera dapat dilakukan melalui *zooming* lensa (Li & Lavest 1996). Proses *zooming* memberikan efek perbesaran atau pengecilan pada citra yang terekam kamera (Hartley & Zisserman 2002). Dalam model *pinhole camera*, proses *zooming* mengakibatkan perubahan pada parameter intrinsik kamera termasuk *focal length* (Willson 1994). Estimasi parameter intrinsik lebih mudah dilakukan dengan menggunakan objek khusus seperti pola papan catur (Zollner & Sablatnig 2004). Namun selama proses *zooming*, mengkalibrasi kamera secara menggunakan objek khusus seperti pola papan catur setiap frame untuk mendapatkan *focal length* merupakan hal yang kurang praktis untuk diterapkan. Untuk mengatasi hal ini proses *self-calibration* perlu dilakukan.

*Self-calibration* adalah proses mendapatkan parameter-parameter intrinsik kamera dengan memanfaatkan kumpulan citra tanpa objek khusus seperti pola papan catur. Proses ini sangat penting untuk mengatasi permasalahan-permasalahan kalibrasi ulang pada kamera yang sering mengalami perubahan *zoom* (Elamsy et al. 2014). *Self-calibration* untuk mendapatkan parameter intrinsik dilakukan pada kamera tunggal (Elamsy et al. 2014, M. Pollefeys & Gool 1998). Dengan demikian, *self-calibration* pada kamera stereo dilakukan pada masing-masing kamera secara terpisah. Hal ini tentunya memerlukan korespondensi titik-titik antar frame pada masing-masing kamera sehingga pada setiap dua frame, pencarian korespondensi untuk proses rekonstruksi pada 4 pasang citra.

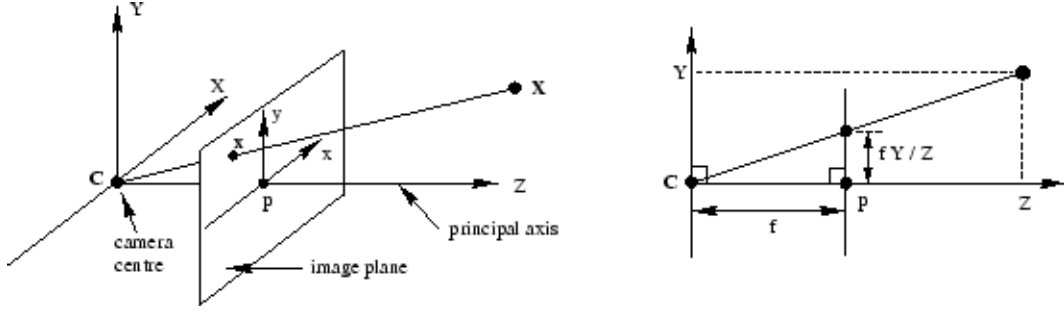
## 2.2. Teori Dasar

Bagian ini membahas singkat teori-teori yang digunakan sebagai dasar dari penelitian yang telah dilakukan.

### 2.2.1. Rekonstruksi Tiga Dimensi Citra

Proses rekonstruksi 3D citra dapat didefinisikan sebagai proses mengolah citra-citra dua dimensi menjadi bentuk 3D yang lebih konstruktif dan membantu pengamat dalam menganalisa informasi-informasi yang terkandung di dalam citra. Rekonstruksi 3D citra sangat penting pada bidang kedokteran, game dan robotika. Secara umum, teknik rekonstruksi terbagi menjadi dua jenis antara lain metode aktif dan pasif. Pada metode aktif, sistem mengukur objek secara aktif baik secara mekanis maupun radiometri. Hal ini dilakukan dengan mengirimkan sinyal pengukur radio (elektromagnetik) atau mekanis (audio) kemudian menganalisa sinyal yang dipantulkan oleh objek. Teknik yang tersebut memerlukan perangkat khusus yang cukup mahal sehingga mempersulit implementasi pada hal-hal baru. Pada metode pasif, sistem menganalisa sinyal-sinyal yang dipantulkan objek tanpa mengirimkan sinyal pengukur. Sistem ini dapat melibatkan penggunaan sensor-sensor yang sensitif terhadap cahaya inframerah atau cahaya tampak seperti pada kamera digital pada umumnya. Metode pasif relatif lebih aplikatif karena perangkat yang digunakan relatif mudah didapat dan ditemukan. Hal ini mendukung penerapan-penerapan baru pada bidang komputer grafis, virtual reality dan komunikasi yang kebutuhannya dalam rekonstruksi 3D meningkat beberapa tahun terakhir (Khilar et al. 2013).

Proses rekonstruksi ini membutuhkan dua atau lebih citra dari sudut pandang pengambilan yang berbeda. Setiap citra memiliki piksel-piksel yang berkorespondensi terhadap citra lain. Piksel-piksel yang saling berkorespondensi antar citra merupakan piksel-piksel hasil proyeksi dari suatu titik pada ruang tiga dimensi yang sama. Pemodelan matematis proyeksi tiga dimensi pada bidang citra biasanya dimodelkan dengan *pinhole camera* seperti pada gambar 2.1. Pada model *pinhole camera*, citra terbentuk dari proyeksi titik tiga dimensi pada bidang citra. Pada gambar 2.1, titik tiga dimensi  $\mathbf{X}$  terproyeksikan menuju pusat proyeksi  $C$  pada bidang citra di titik  $\mathbf{x}$ . Proyeksi titik tiga dimensi  $\mathbf{X}$  pada bidang citra di titik  $\mathbf{x}$  ini direpresentasikan dalam



Gambar 2.1: Model *pinhole camera*

matriks proyeksi  $P$ . Pada sudut pandang atau kamera yang berbeda dengan pusat proyeksi  $C'$ , titik  $\mathbf{X}$  terproyeksi pada bidang citra pada titik  $\mathbf{x}'$ . Pada kondisi ini proyeksi titik  $\mathbf{X}$  pada  $\mathbf{x}'$  direpresentasikan dalam matriks proyeksi  $P'$ . Kedua proses ini dapat dilihat pada persamaan 2.1. Karena titik  $\mathbf{x}$  dan  $\mathbf{x}'$  pada bidang citra merupakan proyeksi dari titik tiga dimensi yang sama, maka kedua titik tersebut dapat dinyatakan titik-titik yang berkorespondensi dan dapat dinyatakan dengan relasi  $\mathbf{x} \leftrightarrow \mathbf{x}'$

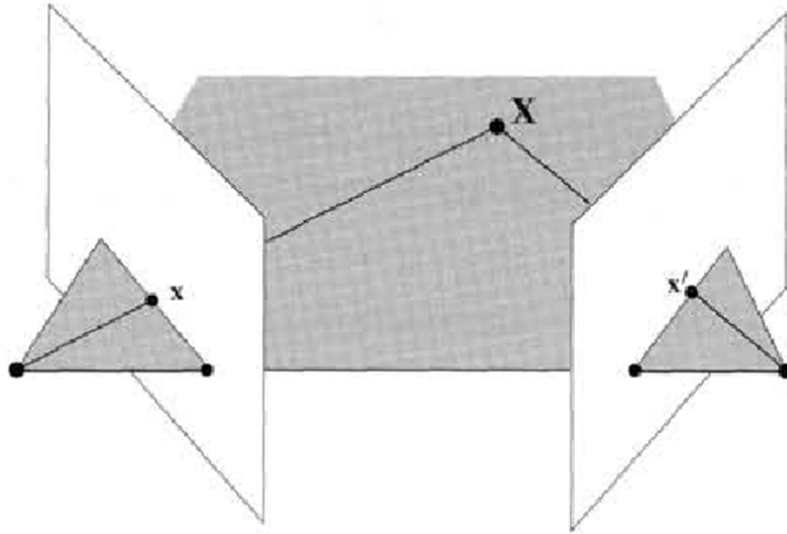
$$\mathbf{x} = P\mathbf{X} \quad \text{dan} \quad \mathbf{x}' = P'\mathbf{X} \quad (2.1)$$

Matriks proyeksi  $P$  bisa didapat dari persamaan persamaan 2.2 di mana  $K$  dan  $K'$  merupakan matriks intrinsik kamera yang mengandung elemen-elemen berupa parameter intrinsik kamera. Matriks  $R$  dan  $t$  merupakan transformasi kamera relatif terhadap koordinat dunia di mana  $R$  merepresentasikan transformasi berupa rotasi dan matriks  $t$  merepresentasikan translasi.

$$P = K[R|t] \quad \text{dan} \quad P' = K'[R'|t'] \quad (2.2)$$

Matriks intrinsik kamera  $K$ , memiliki elemen berupa *focal length*  $f_x, f_y$  dari kamera dan *principal point*  $p_x, p_y$  pada citra. Pada perubahan *zoom* kamera, nilai dari masing-masing *focal length* ikut berubah secara linear.

$$K = \begin{bmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.3)$$



Gambar 2.2: Triangulasi (Hartley & Zisserman 2002)

Dengan diketahuinya matriks kamera  $P$  dan  $P'$ ,  $\mathbf{x}$  dan  $\mathbf{x}'$  adalah dua titik yang memiliki korespondensi, kita dapat mengetahui titik-titik  $C$ ,  $C'$ ,  $\mathbf{x}$  dan  $\mathbf{x}'$  pada sistem koordinat dunia. Dengan demikian, ini berarti bahwa jika kita mengembalikan garis proyeksi dari titik  $\mathbf{x}$  dan  $\mathbf{x}'$  maka garis proyeksi tersebut akan berpotongan pada suatu titik di mana titik perpotongan tersebut dan titik-titik pusat proyeksi dari kedua kamera, titik  $\mathbf{x}$  dan  $\mathbf{x}'$  berada pada satu bidang yang sama yang disebut dengan bidang epipolar. Proses ini disebut dengan triangulasi. Ilustrasi triangulasi dapat dilihat pada gambar 2.2

### 2.2.2. Citra Stereo

Citra stereo merupakan citra yang ditangkap atau dihasilkan oleh dua buah kamera dari sudut pandang yang berbeda. Teknik untuk mendapatkan citra stereo yang memiliki informasi kedalaman (*depth*) disebut dengan *stereoscopy*. Pada era digital ini, *stereoscopy* dikembangkan menyesuaikan teknologi digital dengan kamera digital yang dilengkapi dengan dua atau lebih lensa dengan sensor citra yang terpisah untuk setiap lensa. Hal ini memungkinkan kamera untuk mensimulasikan penglihatan *binocular* manusia dan memberikan informasi tiga dimensi pada citra. Dengan membandingkan dua citra, informasi kedalaman relatif dapat diperoleh, dalam bentuk *disparity*, yang berbanding

terbalik dengan perbedaan jarak ke objek.

Suatu citra adalah proyeksi dua dimensi dari *scene* tiga dimensi. Karena citra merupakan proyeksi dua dimensi, maka sebagian informasi tiga dimensi hilang ketika diproyeksikan pada satu citra. Untuk mendapatkan informasi tiga dimensi suatu *scene* dari satu citra, sebagian fitur atau informasi di dalam *scene*, seperti ada / tidaknya garis, ukuran objek, atau pencahayaan, tekstur, dll. perlu diketahui. Semuanya menjadi lebih sederhana jika citra yang tersedia lebih dari satu dan diambil dari beberapa sudut pandang. Ketika sudut pandang dan citra yang tersedia lebih dari satu, beberapa relasi geometris dapat diturunkan dari titik-titik tiga dimensi dan proyeksi mereka ke masing-masing citra dua dimensi. Relasi ini dapat diekspresikan secara matematis dengan model *pinhole camera* dan *constraint* antara berbagai titik pada masing-masing citra. Jika hanya tersedia dua citra, penelitian ini biasanya memanfaatkan geometri *epipolar*. Secara umum tidak ada perbedaan matematis antara dua citra yang diambil oleh dua kamera secara bersamaan (*stereoscopic*) atau diambil dengan satu kamera pada sudut pandang berbeda dengan waktu berbeda (*optical flow*/pergerakan kamera). Dalam Robotika kedua kasus ini sangat penting. *Stereoscopy* menghasilkan pengetahuan objek dan posisi rintangan sehingga memberikan informasi yang berguna untuk mendapatkan navigasi yang aman dari robot dalam lingkungan apapun. Di sisi lain estimasi pergerakan kamera, dapat dimanfaatkan untuk mendapatkan posisi akhir robot secara spasial.

Secara garis besar metode rekonstruksi dari dua sudut pandang memiliki tiga tahap antara lain :

1. Mendapatkan titik-titik berkorespondensi antar citra
2. Menghitung matriks fundamental dari titik-titik yang berkorespondensi
3. Mencari matriks kamera dari matriks fundamental
4. Menghitung titik 3D dalam *scene* untuk setiap titik yang berkorespondensi.



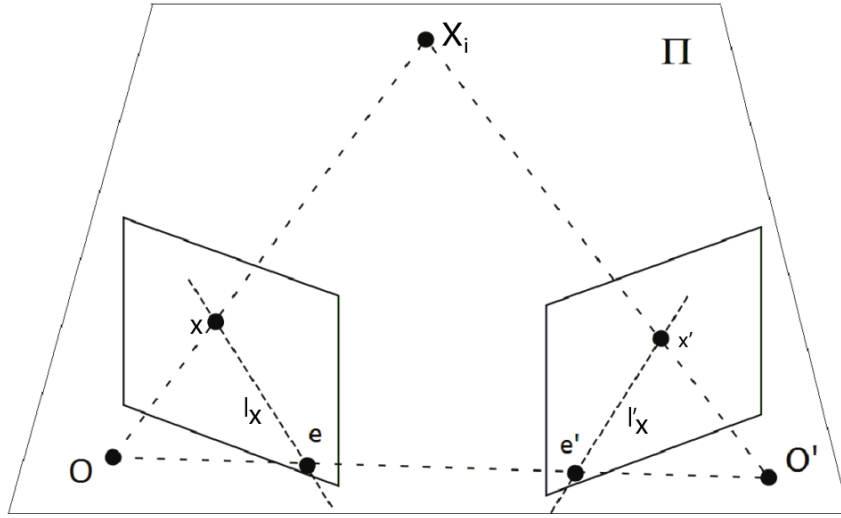
Pencarian titik-titik yang berkorespondensi pada dua citra dapat dibantu dengan memanfaatkan geometri epipolar. Geometri epipolar atau geometri dua citra dan biasanya direpresentasikan dalam bentuk matriks *fundamental*  $3 \times 3$ . Matriks ini memberikan *constraint* untuk titik-titik yang berkorespondensi dan dimungkinkan untuk mengambil semua informasi geometris yang relevan, seperti rotasi dan translasi antar kamera. Estimasi fundamental matriks didasarkan pada kumpulan titik-titik fitur yang tertangkap dan berkorespondensi di kedua citra dari *scene* yang sama. Kesalahan dalam proses menentukan matriks *fundamental* ini secara langsung terkait dengan akurasi dalam perhitungan korespondensi titik-titik pada citra (Bhatti & Navahandi 2011).

Jika diketahui kumpulan titik  $\mathbf{x}_i$  sejumlah  $n$  ( $i=1, 2, \dots, n$ ) pada citra referensi, titik setiap titik  $\mathbf{x}'_i$  yang memenuhi korespondensi  $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$  harus memenuhi suatu *constraint* yang direpresentasikan oleh matriks fundamental  $F$  seperti pada persamaan 2.4

$$\mathbf{x}'_i{}^T F \mathbf{x}_i = 0 \quad \text{untuk semua } i \quad (2.4)$$

Dengan diketahuinya  $\mathbf{x}_i$  dan  $\mathbf{x}'_i$  persamaan ini akan memenuhi syarat untuk beberapa jenis matriks  $F$ . Semakin banyak korespondensi yang diketahui, maka kemungkinan matriks  $F$  yang memenuhi kondisi semakin sedikit dan akurat.

Bayangkan dua citra dari *scene* tiga dimensi yang sama diambil oleh dua kamera di dua sudut pandang yang berbeda (lihat Gambar 2.3). Geometri epipolar mendefinisikan geometri pencitraan dari dua kamera, baik sistem *stereoscopic* atau kamera bergerak tunggal. Diketahui titik tiga dimensi  $\mathbf{X}$  dan proyeksi  $\mathbf{x}$  dan  $\mathbf{x}'$  pada dua bidang citra dari kamera, tiga titik tersebut terletak pada suatu bidang  $\Pi$  yang memotong dua bidang citra pada garis epipolar  $l_x$  dan  $l'_x$  sementara  $e$  dan  $e'$  adalah *epipole*, yaitu titik perpotongan antara pusat proyeksi kamera dengan masing-masing bidang citra. Inti dari *constraint* epipolar hanya menyatakan bahwa jika suatu titik terproyeksikan pada titik  $\mathbf{x}$  pada salah satu bidang citra, maka titik yang berkorespondensi di citra lain-



Gambar 2.3: Geometri Epipolar (Bhatti & Navahandi 2011).

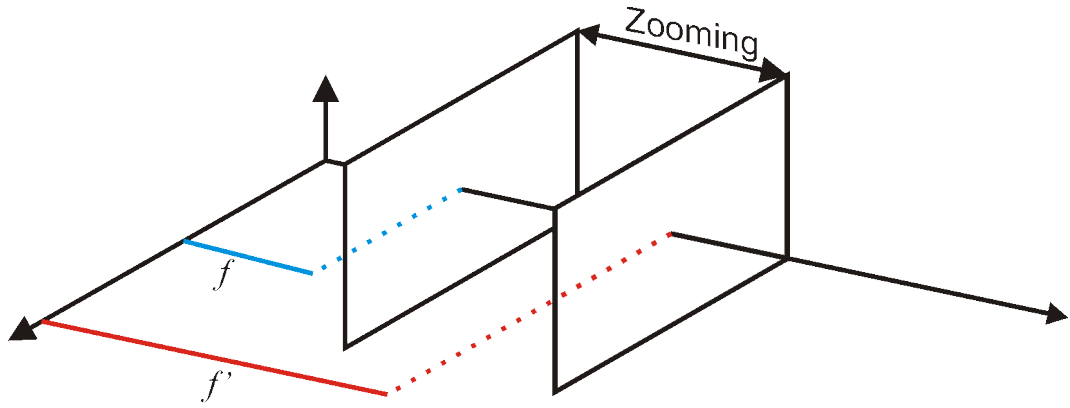
nya harus terletak pada garis epipolar  $l'_x$ . Dari persamaan 2.4, garis *epipolar* dapat didefinisikan seperti pada persamaan 2.5.

$$l'_i = F \mathbf{x}_i \quad \text{untuk semua } i \quad (2.5)$$

Untuk konfigurasi kamera stereo yang memiliki sumbu proyeksi paralel, titik *epipole* ada pada tak hingga sehingga setiap garis epipolar pada citra bersifat paralel antara satu dengan yang lain.

### 2.2.3. *Zooming*

*Zooming* merupakan salah satu metode untuk mendapatkan citra nyata dengan dengan resolusi yang berbeda. Pada pemodelan *pinhole camera* proses *zooming* merupakan penggeseran bidang citra mendekati atau menjauhi pusat proyeksi kamera sehingga memberikan efek perubahan pada *focal length* (Gambar 2.4). Pada kondisi nyata, perubahan *zoom* tidak hanya mengakibatkan perubahan *focal length* namun juga pergeseran atau translasi pada *principal point* (Willson 1994, Hartley & Zisserman 2002). Jika dilihat dari hasil citra yang tertangkap, proses *zooming* mengakibatkan citra yang tertangkap tampak membesar atau mengecil. Pada proses *zooming* pergerakan seluruh titik fitur menjauhi atau mendekati titik pusat *zoom* atau biasa disebut *center of expan-*



Gambar 2.4: *Zooming* pada model *pinhole camera*

*sion* dengan faktor tertentu. Jika suatu titik  $\mathbf{x}$  pada citra bergeser menuju  $\mathbf{x}'$  pada proses *zooming* dengan *center of expansion*  $c$ , maka relasi antara  $\mathbf{x}$  dan  $\mathbf{x}'$  dalam koordinat non-homogen akan memenuhi persamaaan 2.6

$$\mathbf{x}' = s \mathbf{x} + (1 - s)c \quad (2.6)$$

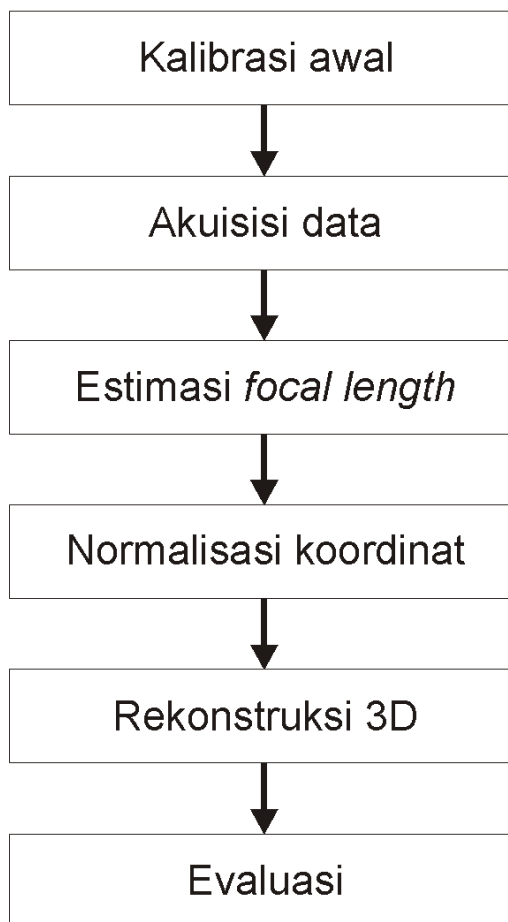
Di mana  $s$  adalah faktor skala atau rasio dari *focal length* sebelum dan sesudah *zooming* (Hartley & Zisserman 2002). Pada kondisi kamera yang diam, titik *center of expansion* ini tidak bergerak selama proses *zooming* (Willson 1994).

*Halaman ini sengaja dikosongkan*

## BAB 3

### METODOLOGI PENELITIAN

Secara garis besar metodologi penelitian dibagi menjadi enam tahap yaitu kalibrasi awal, akuisisi data *zoom*, estimasi *focal length* otomatis, normalisasi koordinat titik, rekonstruksi titik-titik 3D dan evaluasi hasil. Alur metodologi yang digunakan dapat dilihat pada Gambar 3.1.



Gambar 3.1: Metodologi penelitian

#### 3.1. Kalibrasi

Tahap ini diperlukan untuk mengurangi ambiguitas hasil rekonstruksi. Pada tahap ini dilakukan untuk mendapatkan parameter intrinsik dan ekstrinsik



Gambar 3.2: Pola papan catur yang digunakan

kamera. Penelitian ini menggunakan dua jenis kamera yaitu kamera stereo virtual dan kamera stereo fisik. Pada kamera stereo virtual, parameter-parameter kamera ditentukan secara langsung. Pada kamera stereo fisik, kalibrasi kamera awal dilakukan menggunakan pola papan catur seperti kalibrasi pada umumnya dengan ukuran blok  $9 \times 7$ . Kalibrasi dilakukan pada 3 kondisi antara lain *zoom* level minimal, *zoom* level maksimal dan *zoom* level pada *scene* ketiga. Pada tahap ini parameter-parameter kamera didapat dengan menggunakan *Direct Linear Transform* (DLT).

### 3.2. Akuisisi data

Penelitian ini menggunakan dua jenis data untuk rekonstruksi yaitu data sintesis dan data citra kamera. Penggunaan data sintesis dilakukan dengan tujuan mendapatkan *ground truth* titik-titik tiga dimensi serta *focal length* untuk evaluasi pada masing-masing *sequence* yang sulit dilakukan dengan menggunakan kamera fisik. Pada data sintesis, objek yang digunakan model tiga dimensi digital. Model tersebut dipasang pada sistem stereo virtual dengan menyimulasikan dua *pinhole camera* virtual. Data sintesis dihasilkan dengan melakukan simulasi proyeksi model digital terhadap kedua kamera virtual.

Proses *zoom* pada sistem ini dilakukan dengan menentukan *focal length* dari kamera virtual secara langsung pada masing-masing *sequence* proyeksi.

Pengambilan data citra dilakukan dengan merekam empat jenis *scene*. *Scene* pertama dikondisikan terdapat objek yang memiliki fitur-fitur yang dikondisikan (seperti pola papan catur) untuk meminimalkan kesalahan korespondensi titik fitur. *Scene* kedua dan ketiga menyorot pada suatu objek pada lingkungan dalam ruangan dengan titik-titik fitur yang tidak dikondisikan. *Scene* keempat berupa lingkungan luar ruangan tanpa fitur-fitur yang dikondisikan.

### 3.3. Estimasi *focal length* otomatis

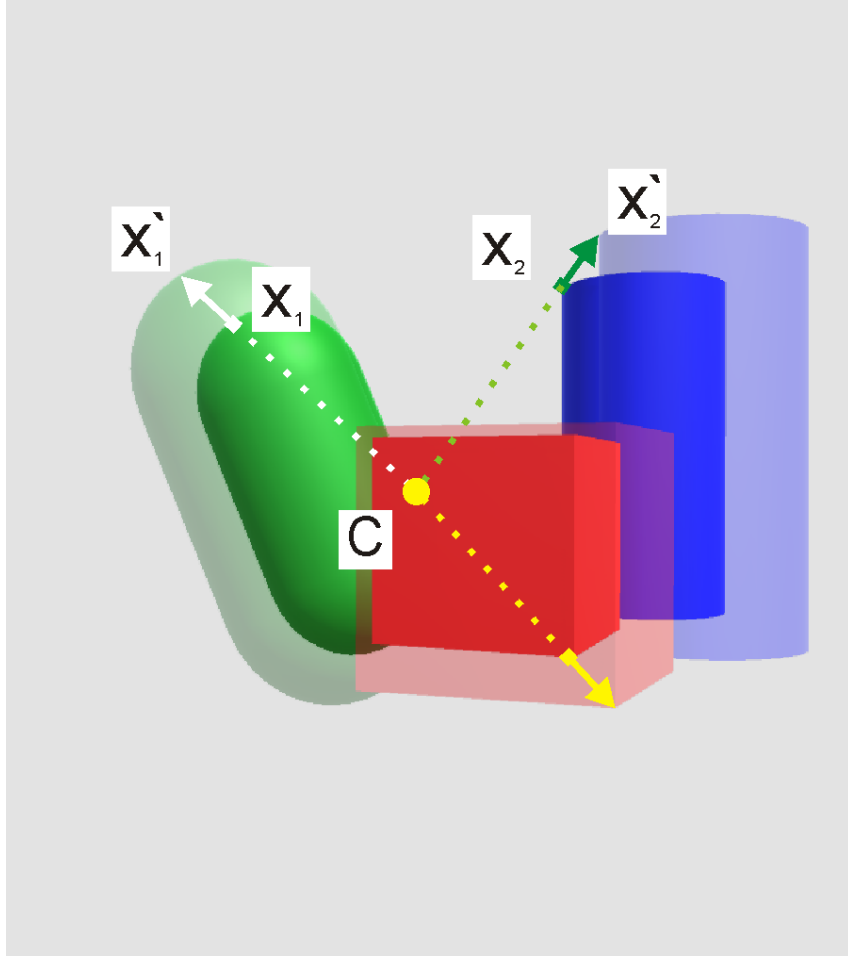
Pada proyeksi, proses *zoom* merupakan proses menggerakkan bidang citra atau proyeksi mendekati atau menjauhi pusat proyeksi (*optical center*). Hal ini tentunya mempengaruhi parameter *focal length* dari kamera. Pada citra yang mengalami perubahan *zoom*, seluruh titik dari citra sebelumnya bergerak mendekat atau menjauh dari *principal plane*. Bagian ini dibagi menjadi dua tahapan yaitu pencarian *focal length* berbasis jarak titik dan pencarian *focal length* menggunakan fundamental matriks. Secara umum estimasi *focal length* diawali dengan mencari faktor skala antar frame. Faktor skala pada frame ke- $n$  ( $s_n$ ) merupakan rasio dari  $f_n$  terhadap  $f_{n-1}$  seperti pada persamaan 3.1.

$$s_n = \frac{f_n}{f_{n-1}} \quad (3.1)$$

Total skala pada frame ke- $n$  ( $S_n$ ) didapat dengan melakukan perkalian kumulatif  $s_i$  dengan nilai  $i = (1, 2 \dots n)$  seperti pada persamaan 3.2.

$$S_n = \prod_{i=1}^n s_i \quad (3.2)$$

Asumsikan nilai *focal length*  $f_0$  adalah *focal length* pada frame pertama, maka nilai  $f_n$  didapatkan melalui persamaan 3.3.



Gambar 3.3: Perubahan posisi titik-titik fitur pada perubahan *zoom*

$$f_n = f_0 S_n \quad (3.3)$$

### 3.3.1. Estimasi *focal length* berbasis jarak titik

Pada tahap ini, beberapa titik fitur yang tertangkap oleh masing-masing kamera digunakan sebagai penentu nilai *focal length*. Anggap kita memiliki titik tiga dimensi  $\mathbf{X}_1$  dan  $\mathbf{X}_2$  diprojektikan oleh kamera dengan *focal length*  $f$  pada titik  $\mathbf{x}_1$  dan  $\mathbf{x}_2$ . Setelah proses *zooming* terjadi, *focal length* berubah menjadi  $f'$  dan mengakibatkan  $\mathbf{X}_1$  dan  $\mathbf{X}_2$  diprojektikan oleh kamera pada titik  $\mathbf{x}'_1$  dan  $\mathbf{x}'_2$ . Berdasarkan persamaan 2.6, kita mendapatkan persamaan 3.4.



$$\begin{aligned}\mathbf{x}'_1 &= s \mathbf{x}_1 + (1-s)c \\ \mathbf{x}'_2 &= s \mathbf{x}_2 + (1-s)c\end{aligned}\tag{3.4}$$

Dari persamaan 3.4, jarak antara titik  $x'_1$  dan  $x'_2$  didapat melalui persamaan 3.5.

$$\begin{aligned}\mathbf{x}'_1 - \mathbf{x}'_2 &= s \mathbf{x}_1 + (1-s)c - (s \mathbf{x}_2 + (1-s)c) \\ \mathbf{x}'_1 - \mathbf{x}'_2 &= s \mathbf{x}_1 - s \mathbf{x}_2 + (1-s)c - (1-s)c \\ \mathbf{x}'_1 - \mathbf{x}'_2 &= s \mathbf{x}_1 - s \mathbf{x}_2 \\ s &= \frac{\mathbf{x}'_1 - \mathbf{x}'_2}{\mathbf{x}_1 - \mathbf{x}_2}\end{aligned}\tag{3.5}$$

Mengingat faktor skala adalah rasio antar *focal length*, relasi antara titik pada citra dengan *focal length* menjadi persamaan 3.6

$$\frac{f'}{f} = \frac{\mathbf{x}'_1 - \mathbf{x}'_2}{\mathbf{x}_1 - \mathbf{x}_2}\tag{3.6}$$

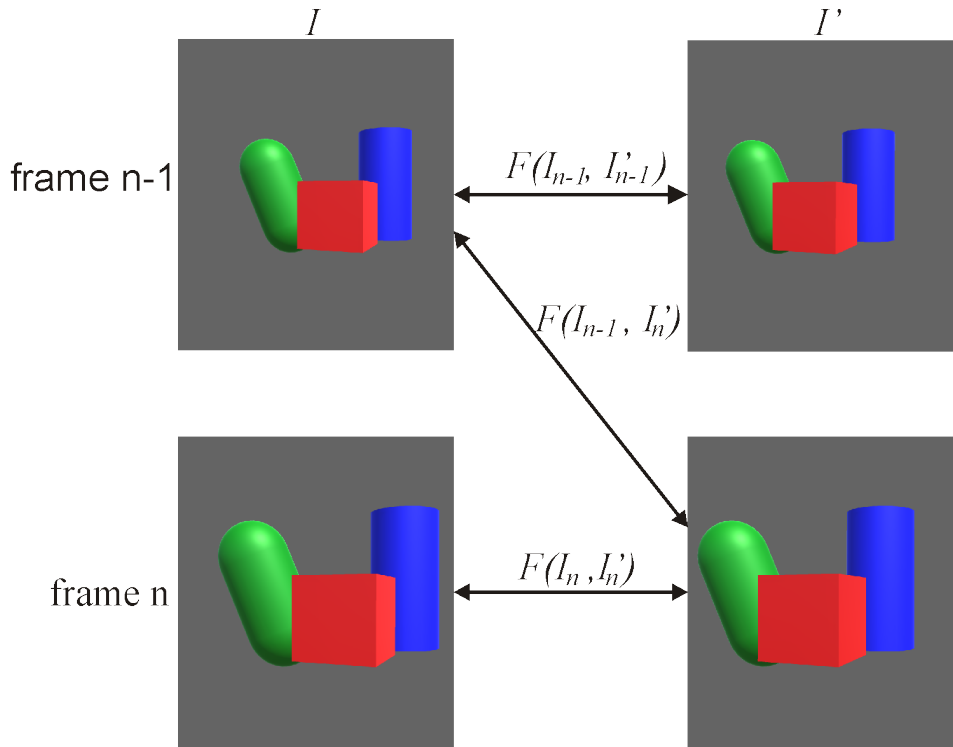
Ilustrasi perubahan zoom dapat dilihat pada Gambar 3.3. Proses pencarian ini dapat dilakukan secara independen antara kamera kiri dan kanan.

Ketika terdapat lebih dari dua pasang titik berkorespondensi ditemukan dalam citra, *Centroid* ( $\bar{\mathbf{x}}$ ) dari distribusi titik pada masing-masing proyeksi digunakan sebagai acuan jarak dari masing-masing titik. Jika terdapat  $n$ -pasang titik berkorespondensi,  $\bar{\mathbf{x}}$  diperoleh melalui persamaaan 3.7

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i\tag{3.7}$$

Nilai  $s$  didapat dari rasio rata-rata jarak  $l^2$  *norm* antara  $\mathbf{x}_i$  centroid  $\bar{\mathbf{x}}$  pada *focal length*  $f'$  terhadap rata-rata jarak pada  $f$ . Nilai  $s$  diperoleh melalui persamaan 3.8.

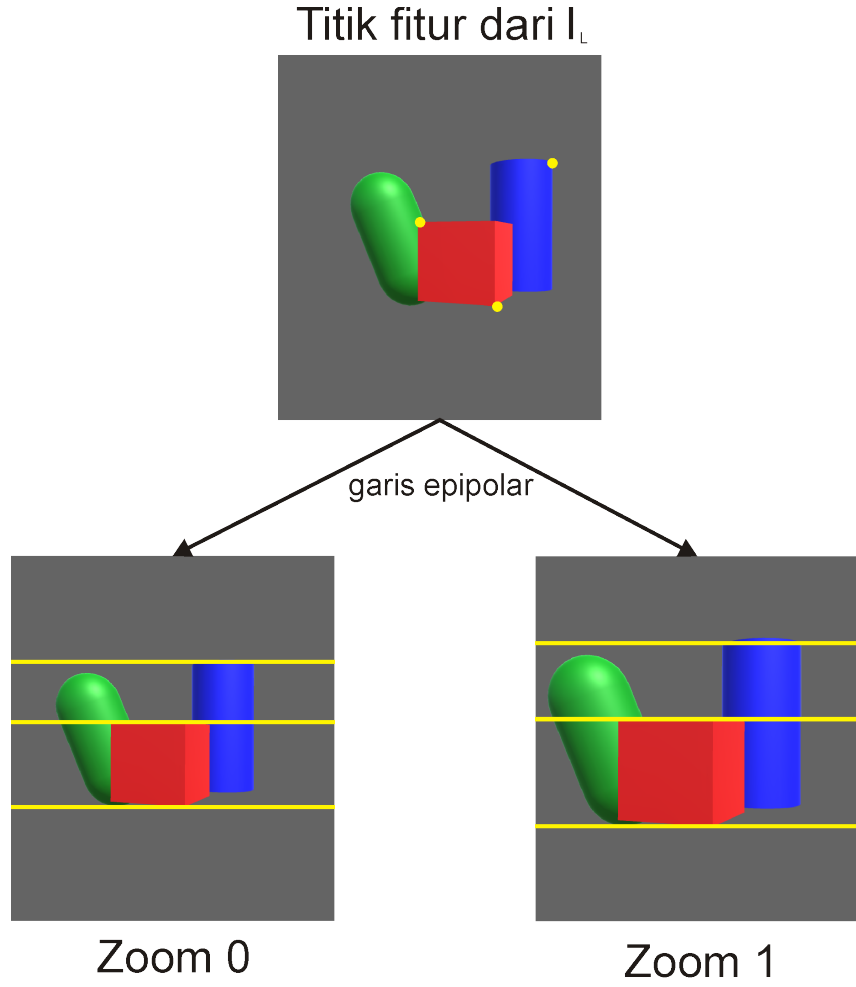
$$s = \frac{\frac{1}{n} \sum_{i=1}^n |\bar{\mathbf{x}}' - \mathbf{x}'_i|}{\frac{1}{n} \sum_{i=1}^n |\bar{\mathbf{x}} - \mathbf{x}_i|}\tag{3.8}$$



Gambar 3.4: Tiga pasangan citra yang digunakan dalam estimasi *focal length* menggunakan fundamental matriks

### 3.3.2. Estimasi *focal length* menggunakan fundamental matriks

Pada tahap ini, beberapa titik fitur dari beberapa pasang citra diambil untuk mendapatkan fundamental matriks sebagai penentu nilai *focal length*. Pada tahap ini, fundamental matriks didapat dari titik-titik fitur yang saling berkorespondensi pada pasangan citra antar kamera dari *sequence* yang sama dan pada pasangan citra dari *sequence* yang berbeda. Ilustrasi pasangan citra yang digunakan dapat dilihat pada 3.4. Tujuan menggunakan fundamental matriks adalah untuk melihat perubahan garis epipolar antar pasangan citra dari frame yang sama maupun dari pasangan citra dari frame yang berbeda. Jika terjadi perubahan *focal length* atau *zoom* maka akan terjadi perubahan *epipolar line* seperti pada 3.5.



Gambar 3.5: Perubahan garis *epipolar* pada *focal length* berbeda

Fundamental matriks didapat dengan asumsi transformasi antar kamera diketahui berupa translasi horizontal murni. Pada kondisi stereo ideal di mana pasangan citra memiliki *principal point* dan *focal length* yang sama, fundamental matriks  $F$  dapat dilihat pada persamaan 3.9.

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.9)$$

Asumsikan kita memiliki pasangan titik berkorespondensi  $\mathbf{x}_i = (x_i, y_i, 1)^T$  dan  $\mathbf{x}'_i = (x'_i, y'_i, 1)^T$ . Dengan menerapkan persamaan 2.5, kita akan mendapatkan garis *epipolar* seperti pada persamaan 3.10. Dengan menerapkan persamaan 2.4, dapat diketahui persamaan garis linear dari garis *epipolar* seperti pada

persamaan 3.11.

$$l'_i = \begin{bmatrix} 0 \\ -1 \\ y_i \end{bmatrix} \quad (3.10)$$

$$0x'_i - y'_i + y_i = 0 \quad \text{atau} \quad y'_i = y_i \quad (3.11)$$

Namun pada kondisi nyata posisi *principal point* dan *focal length* dapat berbeda. Pada kondisi ini, rasio *focal length* memberikan efek *scaling* pada garis *epipolar* dan perbedaan *principal point* secara vertikal mengakibatkan pergeseran garis secara vertikal. Mengetahui ini, persamaan 3.11 dapat diubah menjadi persamaan 3.12 dengan  $a$  sebagai rasio *focal length* dan  $b$  merupakan perbedaan posisi *principal point* secara vertikal. Karena persamaan garis *epipolar* berubah, persamaan fundamental matriks disesuaikan menjadi persamaan 3.13.

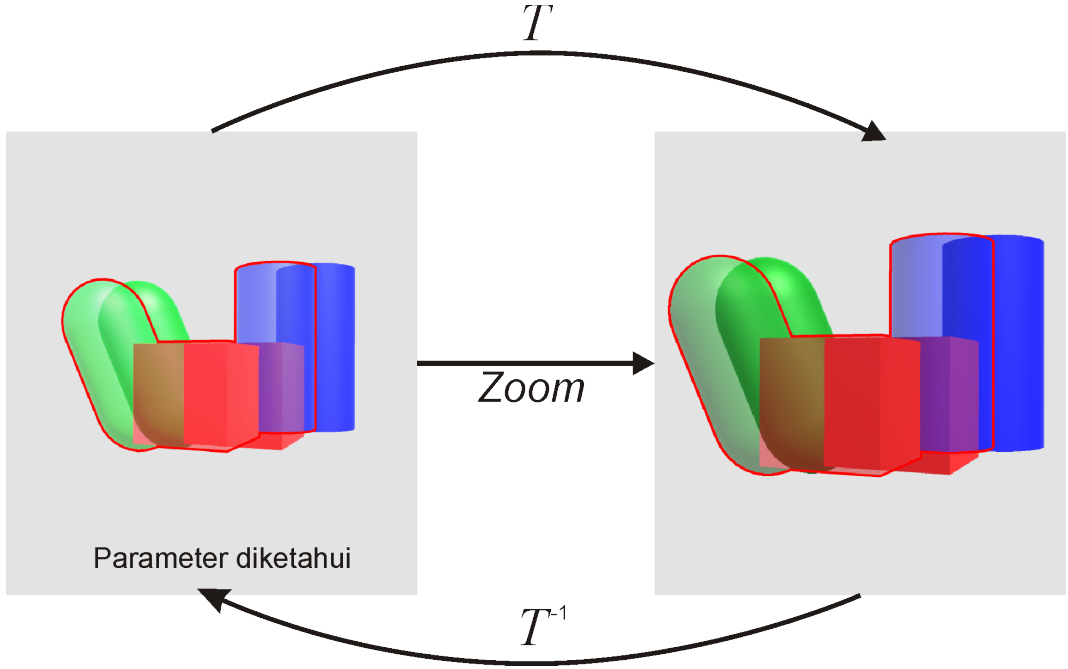
$$y'_i = ay_i + b \quad (3.12)$$

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & a & b \end{bmatrix} \quad (3.13)$$

Misalkan  $a_{u,v}$  dan  $b_{u,v}$  adalah elemen  $a$  dan  $b$  matriks  $F$  pada persamaan 3.13 yang menghubungkan citra  $I_u$  dan  $I'_v$ . Garis epipolar pada citra  $I'_{n-1}$  dan  $I'_n$  yang terhubung pada citra  $I_{n-1}$  dapat diekspresikan melalui persamaan 3.14 dan 3.15

$$y'_{i,n} = a_{n-1,n}y_{i,n-1} + b_{n-1,n-1} \quad (3.14)$$

$$y'_{i,n-1} = a_{n-1,n-1}y_{i,n-1} + b_{n-1,n-1} \quad (3.15)$$



Gambar 3.6: Proses normalisasi menggunakan *inverse* dari *zooming*

Asumsikan  $f_n$  adalah *focal length* dari citra  $I_n$  dan  $f'_n$  adalah *focal length* citra  $I'_n$ , faktor skala  $s$  didapat melalui persamaan 3.16.

$$\frac{a_{n-1,n}}{a_{n-1,n-1}} = \frac{\frac{f'_n}{f_n}}{\frac{f'_{n-1}}{f_n}} = \frac{f'_n}{f'_{n-1}} = s \quad (3.16)$$

Jika persamaan 3.16 dengan persamaan 3.13 digabungkan, estimasi perubahan *focal length* antar frame dapat dilakukan menggunakan persamaan 3.17 di mana  $F_r^c$  merupakan element fundamental matriks  $F$  pada baris ke- $r$  kolom ke- $c$ .

$$s = \frac{F_3^2(I_{n-1}, I'_n)}{F_3^2(I_{n-1}, I'_{n-1})} \quad (3.17)$$

### 3.4. Normalisasi koordinat titik

Normalisasi titik dilakukan karena selama proses *zooming* titik-titik tiga dimensi yang sama terpojeksikan pada koordinat yang berbeda pada setiap perubahan *focal length*. Normalisasi dilakukan dengan menerapkan transfor-

masi pada titik-titik yang mengalami *zooming* menuju sistem koordinat citra yang parameter proyeksinya diketahui (dalam hal ini citra pada frame pertama yang digunakan). Ilustrasi proses normalisasi disajikan pada Gambar 3.6.

Setiap titik yang mengalami *zooming* tranformasi berupa magnifikasi dari titik *expansion center*. Dengan demikian, dalam sistem koordinat homogen terdapat suatu matriks transformasi  $T$  yang menyebabkan titik  $\mathbf{x}_i$  berubah menjadi  $\mathbf{x}'_i$ . Relasi antara  $\mathbf{x}'_i$  dan  $\mathbf{x}_i$  dapat dilihat pada persamaan 3.18.

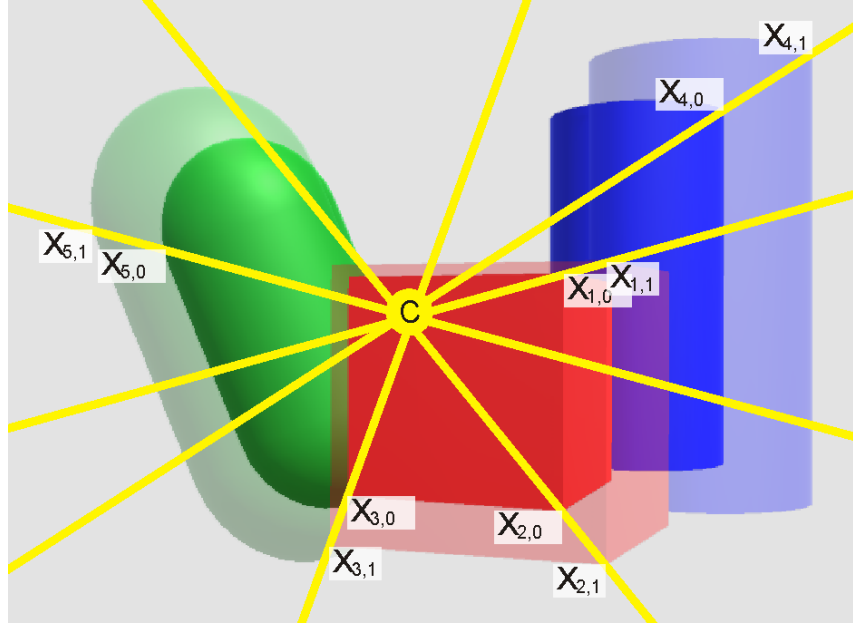
$$\mathbf{x}'_i = T \mathbf{x}_i \quad (3.18)$$

Dalam sistem koordinat non-homogen persamaan 2.6 dapat berubah bentuk menjadi persamaan 3.19.

$$\begin{aligned} \mathbf{x}'_i &= s \mathbf{x}_i + (1 - s)c \\ \mathbf{x}'_i &= s \mathbf{x}_i + c - sc \\ \mathbf{x}'_i &= s(\mathbf{x}_i - c) + c \end{aligned} \quad (3.19)$$

Pada persamaan 3.19  $s$  merupakan faktor skala antar citra dan  $c(x_c, y_c)^T$  merupakan titik *center of expansion*. Persamaan non-homogen 3.19 menunjukkan bahwa titik  $\mathbf{x}_i$  mengalami translasi sejauh  $-c$  dilanjutkan dengan skala sebesar  $s$  dan diakhiri dengan translasi sejauh  $s$ . Dari persamaan 3.19 dapat disimpulkan bahwa matriks transformasi  $T$  pada persamaan 3.18 merepresentasikan gabungan dari transformasi translasi  $\mathbf{t}_c$  dan skala  $\mathbf{s}$ . Berdasarkan persamaan 3.19, dekomposisi matriks  $T$  dapat dilakukan sehingga menjadi persamaan 3.20. Elemen matriks  $\mathbf{t}_c$  dan matriks  $\mathbf{s}$  ditentukan pada persamaan 3.21 dan 3.22.

$$T = \mathbf{t}_c \mathbf{s} \mathbf{t}_c^{-1} \quad (3.20)$$



Gambar 3.7: Garis-garis ekspansi berpotongan pada titik *center of expansion*

$$\mathbf{s} = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.21)$$

$$\mathbf{t}_c = \begin{bmatrix} 1 & 0 & x_c \\ 0 & 1 & y_c \\ 0 & 0 & 1 \end{bmatrix} \quad (3.22)$$

Nilai faktor skala antarframe  $s$  untuk persamaan 3.21 didapat dari tahap estimasi *focal length* namun koordinat dari titik  $c(x_c, y_c)^T$  belum diketahui. Estimasi koordinat *center of expansion* dibagi menjadi dua tahapan yaitu estimasi *center of expansion* berbasis garis ekspansi dan estimasi *center of expansion* menggunakan fundamental matriks.

#### 3.4.1. Estimasi *center of expansion* berbasis garis ekspansi

Pada tahap ini, estimasi posisi  $c(x_c, y_c)$  dilakukan dengan menggunakan perpotongan garis-garis yang dibentuk oleh titik-titik yang mengalami *zooming*. Asumsikan  $\mathbf{x}_{i,0}$  adalah titik-titik pada citra  $I_0$  yang bekorespondensi den-

gan  $x_{i,1}$  pada citra  $I_1$  di mana  $I_1$  adalah  $I_0$  setelah mengalami proses *zooming*. Ilustrasi perpotongan garis ekspansi dapat dilihat pada Gambar 3.7. Dalam sistem koordinat homogen dua dimensi, garis yang melalui dua titik didapatkan melalui *cross product* dari koordinat homogen kedua titik. Dengan demikian garis  $\mathbf{g}_i$  yang melalui titik  $\mathbf{x}_{i,0}$  dan  $\mathbf{x}_{i,1}$  dapat diketahui melalui persamaan 3.23.

$$\mathbf{g}_i = \mathbf{x}_{i,1} \times \mathbf{x}_{i,0} \quad (3.23)$$

Dalam koordinat homogen, perpotongan dua garis dapat diketahui dari hasil *cross product* vektor yang merepresentasikan garis-garis tersebut. Koordinat homogen titik  $\mathbf{c}_{i,j}$  yang merupakan perpotongan antara garis  $g_i$  dan  $g_j$  didapat melalui persamaan 3.24

$$\mathbf{c}_{i,j} = \mathbf{g}_i \times \mathbf{g}_j \quad (3.24)$$

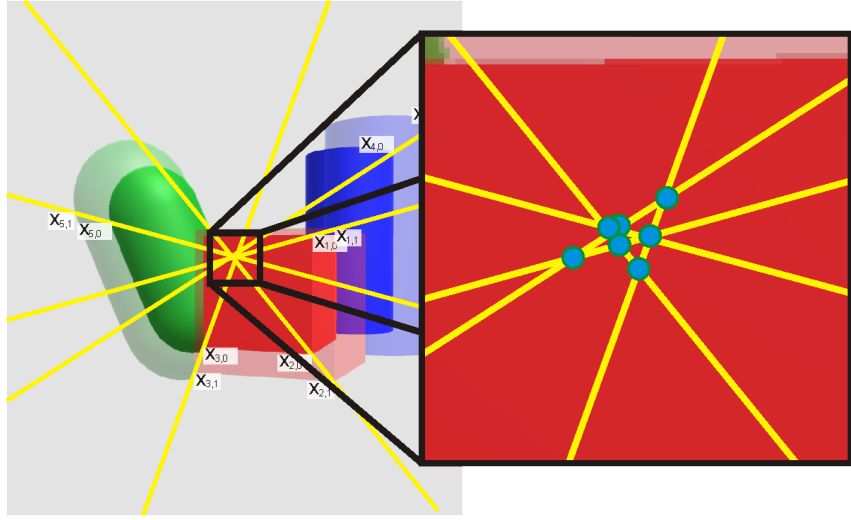
Pada kondisi ideal, semua garis  $\mathbf{g}_i$  akan berpotongan pada titik  $c$ . Namun dalam kondisi pada citra, setiap dua garis dapat berpotongan pada titik-titik yang berbeda-beda seperti pada Gambar 3.8. Perbedaan perpotongan garis dipengaruhi oleh akurasi korespondensi titik yang dapat menurun tergantung dari estimasi posisi pada tingkat sub-piksel. Mengetahui hal ini, koordinat *center of expansion* didapat dengan menggunakan persamaan 3.25

$$c = \frac{1}{m} \sum_{i=1}^m \mathbf{c}_i \quad (3.25)$$

### 3.4.2. Estimasi *center of expansion* menggunakan fundamental matriks

Tahap ini memperhatikan perubahan garis-garis epipolar sebelum dan sesudah zoom terjadi seperti pada gambar 3.5. Asumsikan  $\mathbf{x}_{i,0}$  adalah titik-titik pada citra  $I_0$  yang berkorespondensi dengan  $\mathbf{x}'_{i,0}$  pada citra  $I'_0$  dan  $x'_{i,1}$  pada citra  $I'_1$  di mana  $I_0$  dan  $I'_0$  adalah pasangan citra stereo serta  $I'_0$  adalah  $I'_1$  sete-





Gambar 3.8: Perpotongan garis-garis ekspansi

lah mengalami proses *zooming*. Dalam proses *zooming*, titik-titik pada citra  $I'_0$  dan  $I^{prime}_1$  yang berkorespondensi dengan titik-titik pada citra  $I_0$  memiliki perbedaan. Hal ini juga mengakibatkan garis-garis epipolar pada  $I'_0$  dan  $I'_0$  juga berbeda. Pada persamaan 3.18 diketahui bahwa titik-titik pada citra  $I^{prime}_0$  mengalami transformasi  $T$  untuk mencapai posisi  $I^{prime}_n$ . Dengan demikian, dapat disimpulkan bahwa terdapat suatu matriks transformasi  $T_l$  yang mentransformasi garis-garis epipolar pada citra  $I'_0$  ( $l_0$ ) menjadi garis-garis epipolar pada citra  $I_1$  ( $l_1$ ) seperti pada persamaan 3.26.

$$l_1 = T_l l_0 \quad (3.26)$$

Matriks  $T_l$  dipengaruhi oleh matriks  $T$ . Untuk mengetahui relasi antara  $T_l$  dan  $T$ , persamaan 2.4 dan 3.18 digabungkan sehingga mendapatkan persamaan 3.27.

$$\mathbf{x}'_{i,0}{}^T F(I_0, I'_0) \mathbf{x}_{i,0} = 0 \quad \text{dan} \quad (T \mathbf{x}'_{i,0})^T F(I_0, I'_1) \mathbf{x}_{i,0} = 0 \quad (3.27)$$

Dengan menggunakan substitusi pada persamaan 2.5, persamaan *constraint* epipolar didapat seperti pada persamaan 3.28.

$$\mathbf{x}_{i,0}'^T l'_0 = 0 \quad \text{dan} \quad (T \mathbf{x}_{i,0}')^T l'_1 = 0 \quad (3.28)$$

Persamaan 3.28 membuktikan bahwa matriks  $T$  mempengaruhi transformasi dari garis  $l_0$  menjadi  $l_1$ . Melanjutkan persamaan 3.28,  $T_l$  didapatkan melalui persamaan 3.29

$$\begin{aligned} \mathbf{x}_{i,0}'^T l'_0 &= 0 & \mathbf{x}_{i,0}'^T T^T l'_1 &= 0 \\ \mathbf{x}_{i,0}'^T (l'_0) &= 0 & \mathbf{x}_{i,0}'^T (T^T l'_1) &= 0 \\ T^T l'_1 &= l'_0 \\ l'_1 &= (T^T)^{-1} l' \\ \text{sehingga} \\ T_l &= (T^T)^{-1} \end{aligned} \quad (3.29)$$

Dengan memperhatikan persamaan 3.20, 3.21 dan 3.22, Kita mendapatkan elemen-elemen dari matriks  $T_l$  melalui persamaan 3.30.

$$\begin{aligned} T_l &= (\mathbf{t} \mathbf{s} \mathbf{t}^{-1})^{-T} \\ T_l &= \left( \begin{bmatrix} 1 & 0 & x_c \\ 0 & 1 & y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & 1 \end{bmatrix} \right)^{-T} \\ T_l &= \begin{bmatrix} s & 0 & x_c - s(x_c) \\ 0 & s & y_c - s(y_c) \\ 0 & 0 & 1 \end{bmatrix}^{-T} \\ T_l &= \begin{bmatrix} \frac{1}{s} & 0 & 0 \\ 0 & \frac{1}{s} & 0 \\ \frac{s(x_c) - x_c}{s} & \frac{s(y_c) - y_c}{s} & 1 \end{bmatrix} \end{aligned} \quad (3.30)$$

Jika kita substitusi nilai  $l_0$  dan  $l_1$  pada persamaan transformasi garis 3.26 dengan garis  $l_0$  dan  $l_1$  menggunakan persamaan 2.5, kita akan mendapatkan persamaan 3.31.

$$\begin{aligned}
l_1 &= T_l l_0 \\
F(I_0, I'_1) \mathbf{x}_{i,0} &= T_l F(I_0, I'_0) \mathbf{x}_{i,0}
\end{aligned} \tag{3.31}$$

Jika kita mengoperasikan persamaan 3.31 dengan melihat elemen-elemen matriks di dalamnya, kita akan mendapatkan dua bentuk garis untuk  $l_1$  seperti pada persamaan 3.32.

$$\begin{aligned}
\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & a_1 & b_1 \end{bmatrix} \begin{bmatrix} x_{i,0} \\ y_{i,0} \\ 1 \end{bmatrix} &= \begin{bmatrix} \frac{1}{s} & 0 & 0 \\ 0 & \frac{1}{s} & 0 \\ \frac{s(x_c) - x_c}{s} & \frac{s(y_c) - y_c}{s} & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & a_0 & b_0 \end{bmatrix} \begin{bmatrix} x_{i,0} \\ y_{i,0} \\ 1 \end{bmatrix} \\
\begin{bmatrix} 0 \\ -1 \\ a_1 y_0 + b_1 \end{bmatrix} &= \begin{bmatrix} \frac{1}{s} & 0 & 0 \\ 0 & \frac{1}{s} & 0 \\ \frac{s(x_c) - x_c}{s} & \frac{s(y_c) - y_c}{s} & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ a_0 y_{i,0} + b_0 \end{bmatrix} \\
\begin{bmatrix} 0 \\ -1 \\ a_1 y_0 + b_1 \end{bmatrix} &= \begin{bmatrix} 0 \\ \frac{-1}{s} \\ a_0 y_{i,0} + b_0 + \frac{y_c - s(y_c)}{s} \end{bmatrix}
\end{aligned} \tag{3.32}$$

Dari persamaan 3.31, dua bentuk persamaan untuk garis  $l_1$  dapat diketahui. Persamaan pertama di dapat dari *epipolar constraint* seperti pada persamaan 3.33 sedangkan persamaan kedua pada 3.34 terbentuk dari transformasi garis epipolar pada citra  $I'_0$ .

$$\begin{aligned}
0x_0 x'_{i,1} - y'_{i,1} + a_1 y_{i,0} + b_1 &= 0 \\
y'_{i,1} &= a_1 y_{i,0} + b_1
\end{aligned} \tag{3.33}$$

$$\begin{aligned}
0x_0x'_{i,1} - \frac{y'_{i,1}}{s} + a_0y_{i,0} + b_0 + \frac{y_c - s(y_c)}{s} &= 0 \\
y'_{i,1} &= s\left(a_0y_{i,0} + b_0 + \frac{y_c - s(y_c)}{s}\right) \\
y'_{i,1} &= s(a_0y_{i,0}) + s(b_0) + y_c - s(y_c)
\end{aligned} \tag{3.34}$$

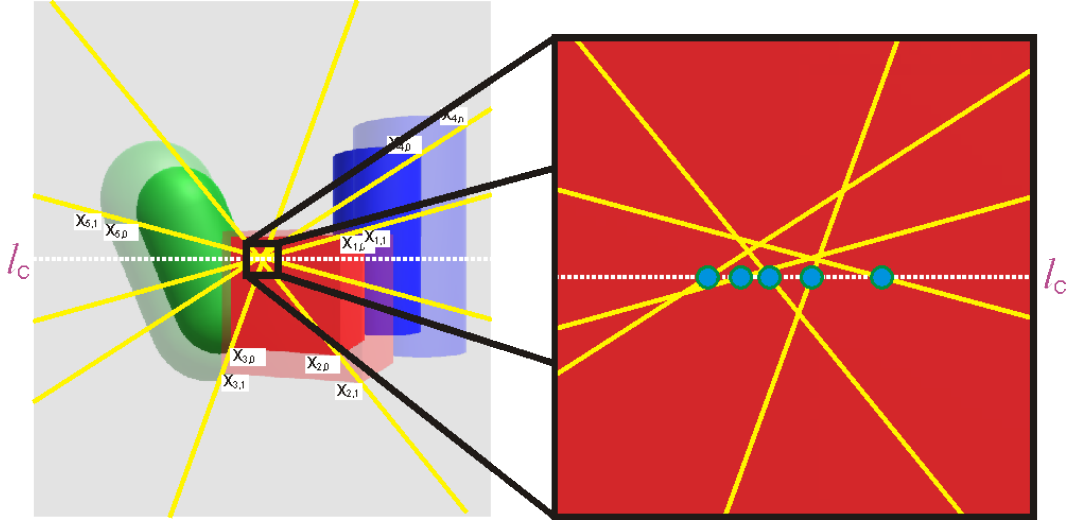
Karena persamaan 3.33 dan 3.34 merupakan persamaan garis yang sama dan dengan melihat kembali persamaan 3.16, kedua persamaan tersebut digabungkan untuk mendapatkan  $y_c$  seperti pada persamaan 3.35

$$\begin{aligned}
s(a_0y_{i,0}) + s(b_0) + y_c - s(y_c) &= a_1y_{i,0} + b_1 \\
y_c - s(y_c) &= a_1y_{i,0} + b_1 - s(a_0y_{i,0}) - s(b_0) \\
y_c(1 - s) &= a_1y_{i,0} + b_1 - \frac{a_1}{a_0}(a_0y_{i,0}) - s(b_0) \\
y_c(1 - s) &= a_1y_{i,0} + b_1 - a_1y_{i,0} - s(b_0) \\
y_c(1 - s) &= b_1 - sb_0 \\
y_c &= \frac{b_1 - s(b_0)}{1 - s}
\end{aligned} \tag{3.35}$$

Setelah mendapatkan  $y_c$ ,  $x_c$  diperoleh dengan mencari titik perpotongan  $\mathbf{c}_i$  antara garis ekspansi  $\mathbf{g}_i$  pada persamaan 3.23 dengan garis  $l_c(0 - 1y_c)^T$  seperti pada persamaan 3.36.

$$\mathbf{c}_i = l_c \times g_i \tag{3.36}$$

Pada kondisi ideal, semua garis  $\mathbf{g}_i$  akan memotong  $\mathbf{l}_c$  pada titik  $c$ . Namun dalam kondisi pada citra, titik perpotongan dapat memiliki koordinat  $x$  yang berbeda seperti pada Gambar 3.9. Perbedaan perpotongan garis dipengaruhi oleh akurasi korespondensi titik yang dapat menurun tergantung dari estimasi posisi pada tingkat sub-piksel. Mengetahui hal ini, persamaan 3.37 digunakan untuk mendapatkan  $c_x$ .



Gambar 3.9: Perpotongan garis-garis ekspansi dengan garis  $l_c$

$$c_x = \frac{1}{m} \sum_{i=1}^m x_{\mathbf{c},i} \quad (3.37)$$

Dengan demikian, semua elemen dari matriks  $T$  telah diketahui dan proses normalisasi dari koordinat titik yang mengalami *zooming*  $\mathbf{x}'_{i,1}$  menuju titik ternormalisasi  $\widetilde{\mathbf{x}'_{i,1}}$  dapat dilakukan menggunakan persamaan 3.38.

$$\widetilde{\mathbf{x}'_{i,1}} = T^{-1} \mathbf{x}'_{i,1} \quad (3.38)$$

### 3.5. Rekonstruksi Tiga Dimensi

Pada tahap ini dilakukan proses rekonstruksi dengan asumsi kamera terkali-brasi. Pada asumsi kamera terkalibrasi, penelitian ini menggunakan hasil dari tahap kalibrasi sebagai parameter kamera dan hasil dari tahap normalisasi sebagai titik-titik citra yang direkonstruksi. Dengan mengetahui parameter intrinsik dalam bentuk matriks  $K$  dan  $K'$  serta parameter ekstrinsik antar kamera dalam bentuk matriks  $R$  dan  $t$ , rekonstruksi menggunakan matriks proyeksi  $P$  dan  $P'$  yang dideskripsikan pada persamaan 3.39.

$$P = [K|0] \quad \text{dan} \quad P' = K'[R, t] \quad (3.39)$$

Setelah mendapatkan  $P$  dan  $P'$ , titik-titik tiga dimensi didapatkan dengan melakukan triangulasi linear. Untuk mendapatkan nilai  $X_i$  kedua persamaan pada 2.1 digabungkan untuk membangun persamaan linear  $AX = 0$ . Mengingat  $x$  dan  $x'$  adalah koordinat homogen, faktor skala homogen dapat dieliminasi dengan melakukan *cross product* seperti pada persamaan 3.40.

$$x_i \times PX_i = 0 \quad \text{dan} \quad x'_i \times PX'_i = 0 \quad (3.40)$$

$$A = \begin{bmatrix} x_{i,1}p^{3T} - p^{1T} \\ x_{i,2}p^{3T} - p^{2T} \\ x'_{i,1}p^{3T} - p^{1T} \\ x'_{i,2}p^{3T} - p^{2T} \end{bmatrix} \quad (3.41)$$

Asumsikan  $p^r$  adalah matriks 1 *times* 3 berisi elemen-elemen dari  $P$  pada baris ke- $r$  dan  $p'^r$  berisi elemen-elemen dari  $P'$  pada baris ke- $r$ , dari persamaan 3.40 elemen-elemen  $A$  didapat menggunakan persamaan 3.41. Nilai  $X$  didapat dari hasil *singular value decomposition* dari matriks  $A$ .

### 3.6. Evaluasi hasil

Evaluasi dilakukan dengan membandingkan rekonstruksi yang melakukan estimasi *focal length* dan *center of expansion* menggunakan menggunakan *fundamental matriks* terhadap estimasi yang menggunakan jarak antar titik dan perpotongan garis-garis ekspansi. Evaluasi ini membedakan metode evaluasi antara data sintetis dan data citra karena data sintetis memiliki *ground truth* yang diketahui.

Evaluasi hasil rekonstruksi dari data sintetis dilakukan dengan menggunakan *mean squared error* (MSE) dan *Pearson correlation coefficient* terhadap model virtual tiga dimensi yang menjadi *ground truth*. Evaluasi MSE dan Pearson dilakukan pada hasil rekonstruksi masing-masing frame sehingga perubahan skor evaluasi seiring dengan perubahan *focal length* dapat terlihat.

Pada data citra hasil rekonstruksi dievaluasi dengan mengumpulkan hasil

rekonstruksi dari kedua metode pada satu ruang. Hasil rekonstruksi tersebut diproyeksikan secara ortogonal ke bidang sumbu  $xy$  dan  $zy$ . Pengumpulan objek dan proyeksi ortogonal dilakukan untuk membandingkan distribusi titik hasil rekonstruksi dari kedua metode. Hasil rekonstruksi dengan distribusi titik tiga dimensi yang lebih kecil menunjukkan bahwa hasil rekonstruksi lebih akurat karena data korespondensi yang digunakan pada kedua metode sama.

*Halaman ini sengaja dikosongkan*



## BAB 4

### HASIL DAN PEMBAHASAN

Bagian ini memaparkan proses-proses yang telah dilakukan dimulai dari akuisisi data hingga hasil rekonstruksi.

#### 4.1. Kalibrasi

Bagian ini menyajikan spesifikasi kamera stereo virtual dan fisik secara terpisah. Tahap ini digunakan untuk mengurangi ambiguitas hasil rekonstruksi. Tahap ini berupaya mendapatkan parameter intrinsik dan ekstrinsik kamera. Penelitian ini menggunakan dua jenis kamera yaitu kamera stereo virtual dan kamera stereo fisik.

##### 4.1.1. Spesifikasi kamera virtual

Pada simulasi rekonstruksi menggunakan data sintetis, penelitian ini menggunakan tiga jenis kamera stereo virtual. Kamera pertama menggunakan kamera stereo virtual dengan parameter intrinsik yang identik serta memiliki nilai *focal length* horizontal ( $f_x$ ) yang sama dengan *focal length* vertikal ( $f_y$ ). Spesifikasi kamera virtual ini dapat dilihat pada Tabel 4.1.

Kamera kedua menggunakan kamera stereo virtual dengan parameter intrinsik yang identik namun memiliki nilai *focal length* horizontal ( $f_x$ ) yang berbeda dengan *focal length* vertikal ( $f_y$ ). Spesifikasi kamera virtual ini dapat dilihat pada Tabel 4.2.

Kamera ketiga menggunakan kamera stereo virtual dengan parameter intrinsik yang berbeda dengan nilai *focal length* horizontal ( $f_x$ ) dan *focal length*

Tabel 4.1: Parameter intrinsik pada kamera stereo virtual pertama

Parameter	Kamera Kiri	Kamera Kanan
$f_x$	500	500
$f_y$	500	500
<i>Image Center</i>	X : 512 Y : 384	X : 512 Y : 384

Tabel 4.2: Parameter intrinsik pada kamera stereo virtual kedua

Parameter	Kamera Kiri	Kamera Kanan
$f_x$	500	500
$f_y$	495	495
<i>Image Center</i>	X : 512 Y : 384	X : 512 Y : 384

Tabel 4.3: Parameter intrinsik pada kamera stereo virtual ketiga

Parameter	Kamera Kiri	Kamera Kanan
$f_x$	500	503
$f_y$	500	493
<i>Image Center</i>	X : 512 Y : 384	X : 510 Y : 390

vertikal ( $f_y$ ) yang berbeda pula. Spesifikasi kamera virtual ini dapat dilihat pada Tabel 4.3.

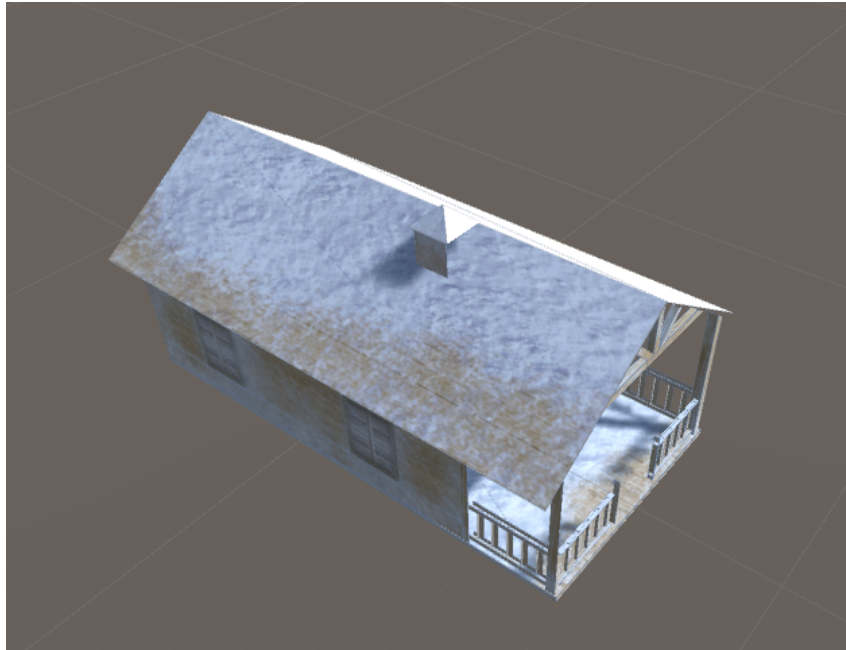
#### 4.1.2. Spesifikasi kamera fisik

Perangkat kamera fisik yang digunakan dalam penelitian ini adalah kamera stereo Sony HDR-TD30VE. Citra yang tertangkap berukuran  $1920 \times 1080$  piksel. Pada kamera stereo fisik kalibrasi dilakukan pada kamera dengan *zoom* minimum. Pada *zoom* minimum, kalibrasi dilakukan dengan menggunakan pola papan catur dengan ukuran persegi  $4 \text{ cm} \times 4 \text{ cm}$ . Hasil kalibrasi pada *zoom* minimal dapat dilihat pada Tabel 4.4.

Pada Tabel 4.4 dapat dilihat bahwa terdapat sedikit perbedaan skala antara sumbu horizontal dan sumbu vertikal pada citra. Hal ini menunjukkan bahwa satu piksel pada citra secara fisik memiliki ukuran lebar dan tinggi yang berbeda. Hal lain yang dapat diamati dari Tabel 4.4 adalah perbedaan *focal*

Tabel 4.4: Parameter intrinsik pada *zoom* minimal

Parameter	Kamera Kiri	Kamera Kanan
$f_x$	1855.8880481487653923977	1855.4648296069308344158
$f_y$	1851.6277503159415118716	1849.3833629084058429725
<i>Image center</i>	X : 953.04581250733201613 Y : 529.33307140597264606	X : 974.71447153449196321 Y : 558.32439405598358916



Gambar 4.1: Model virtual tiga dimensi

*length* serta perbedaan posisi *principal point* antara kamera kiri dan kamera kanan. Hal ini menunjukkan bahwa citra yang ditangkap bukanlah citra stereo ideal. Pada *zoom* maximum, kalibrasi dilakukan menggunakan pola papan catur dengan ukuran persegi  $1 \text{ cm} \times 1 \text{ cm}$ .

## 4.2. Akuisi data

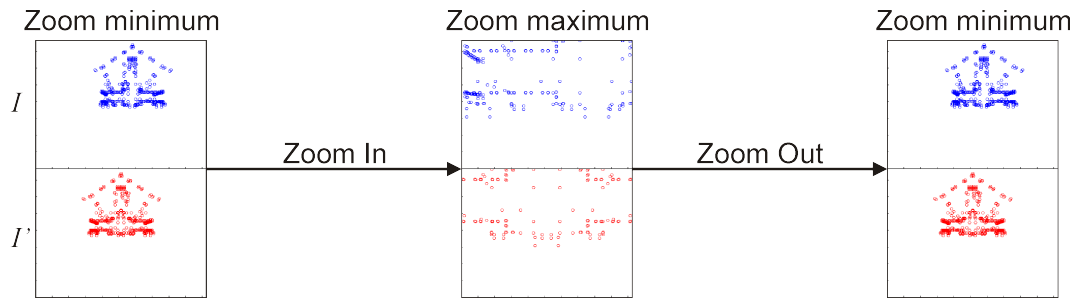
Sesuai dengan yang telah dipaparkan pada subbab 3.2, akuisisi data dilakukan menggunakan dua jenis data koordinat yaitu data sintetis dan data citra. Bagian ini memaparkan proses pengambilan dari kedua jenis data tersebut.

### 4.2.1. Akuisisi data sintetis

Pada pembuatan data sintetis, objek tiga dimensi yang digunakan adalah model objek tiga dimensi virtual berbentuk rumah seperti pada 4.1. Model ini diproyeksikan pada kamera-kamera stereo virtual dengan spesifikasi awal pada Tabel 4.1 hingga Tabel 4.3 dengan simulasi ukuran citra yang sama yaitu *1024 times 768*. Semua kamera virtual akan menyimulasikan proses *zooming* dengan

Tabel 4.5: Spesifikasi simulasi proses *zooming* pada data sintetis

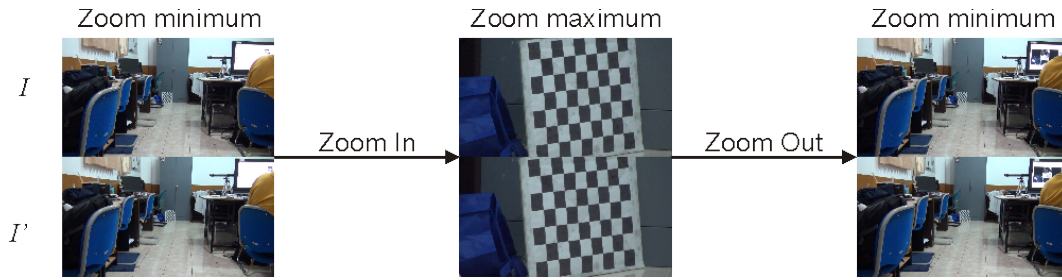
Parameter		Nilai
Faktor skala per frame	Zoom in	1.015×
	Zoom out	0.985221675×
Zoom Level	Min	1×
	Max	4.432×
Jumlah Frame	Idle	30
	Zoom in	100
	Zoom out	100
	Total	230
Ukuran Citra	Lebar	1024 px
	Tinggi	768 px
Simulasi Akurasi	Level 0	Tipe data double
	Level 1	0.01 px
	Level 2	0.1 px
	Level 3	1 px



Gambar 4.2: Hasil proyeksi model tiga dimensi virtual.

perubahan *focal length* dan jumlah serta komposisi frame yang sama. Proses *zoom in* dilakukan dengan faktor skala  $s$  sebesar 1.015 per frame dan untuk *zoom out* sebesar 0.985221675 dengan level *zoom* maksimal sebesar 4.432 kali. Simulasi dilakukan pada 4 tingkat ketelitian proyeksi antara lain : akurasi penuh, akurasi 0.01 piksel, akurasi 0.1 piksel dan akurasi 1 piksel. Detail dari skenario proyeksi dapat dilihat pada Tabel 4.5.

Untuk kamera ketiga akuisisi data dilakukan dengan menggunakan dua skenario proyeksi yaitu proyeksi dengan *image center* yang tetap dan proyeksi dengan *image center* yang dinamis pada setiap perubahan frame yang terjadi proses *zooming*. Sebagian dari hasil proses proyeksi dapat dilihat pada Gambar 4.2.

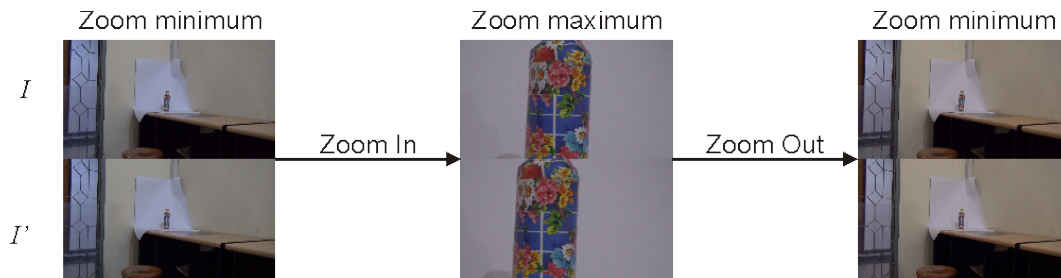
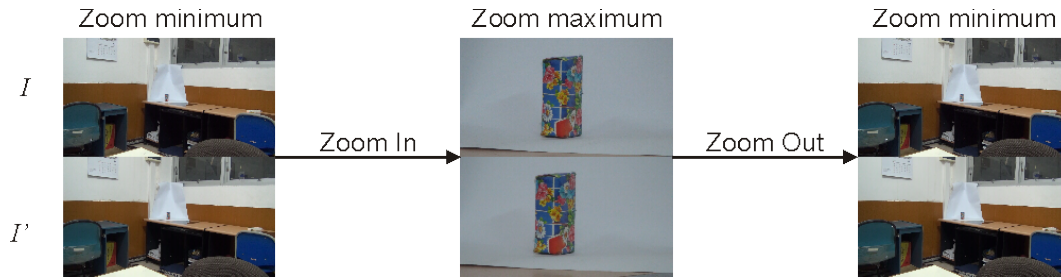


Gambar 4.3: Data *scene* pertama, objek dengan pola papan catur.

#### 4.2.2. Akuisisi data citra

Pengambilan *scene* data citra dilakukan dengan merekam *scene* diawali dengan *zoom* minimal kemudian *zoom in* hingga mencapai maksimal dan dilanjutkan dengan *zoom out* hingga kembali pada *zoom* minimal. Proses *zooming* dihentikan sejenak saat mencapai *zoom* maksimal maupun *zoom* minimal untuk mempermudah identifikasi kondisi *focal length* tertinggi maupun terendah. Pada jenis data citra, data diambil dari empat jenis *scene* yaitu *scene* pertama yang mengandung pola papan catur, *scene* kedua dan ketiga menyorot objek tanpa pola papan catur dan *scene* keempat yang diambil di luar ruangan. Pada *scene* yang mengandung pola papan catur, akuisisi data dilakukan meletakkan pola papan catur pada sekitar 6 meter dari kamera. Hal ini dilakukan karena *zoom* maksimum dari kamera yang digunakan cukup besar sehingga ketika mencapai *zoom* maksimum, tidak ada titik dari pola papan catur yang tidak tertangkap citra. Pola papan catur yang digunakan terdiri atas  $9 \times 7$  persegi dengan masing-masing persegi berukuran  $4 \text{ cm} \times 4 \text{ cm}$ . Ilustrasi pengambilan data dapat dilihat pada 4.3. Data yang didapat sejumlah 702 frame.

Pada *scene* kedua, objek yang disorot berupa objek balok yang diletakkan di dalam ruangan. Perekaman *scene* diawali dengan *zoom* minimal kemudian *zoom in* hingga mencapai maksimal dan dilanjutkan dengan *zoom out* hingga kembali pada *zoom* minimal. Penataan *scene* dilakukan dengan mengkondisikan seluruh bagian dari objek tetap tertangkap kamera pada saat kamera mencapai *zoom* maksimal. Ilustrasi pengambilan data dapat dilihat pada 4.4. Data yang didapat untuk *scene* ini sejumlah 625 frame.

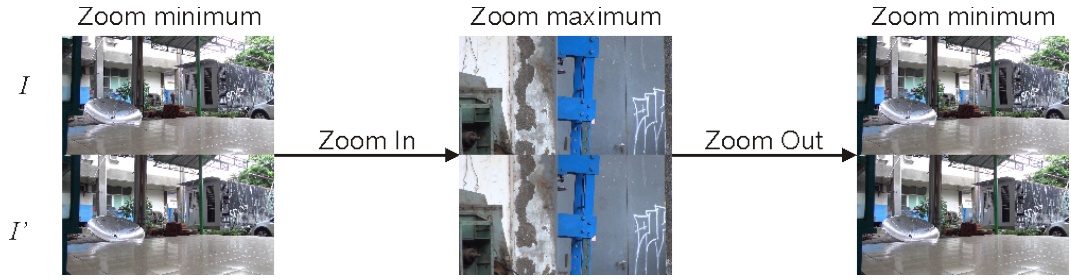


Pada *scene* ketiga, objek yang disorot berupa botol yang diletakkan di dalam ruangan. Perekaman *scene* diawali dengan *zoom* minimal kemudian *zoom in* hingga mencapai maksimal dan dilanjutkan dengan *zoom out* hingga kembali pada *zoom* minimal. Penataan *scene* dilakukan dengan mengkon-disikan sebagian dari objek tidak tertangkap kamera pada saat kamera menca-pai *zoom* maksimal. Ilustrasi pengambilan data dapat dilihat pada 4.5. Data yang didapat untuk *scene* ini sejumlah 592 frame.

Pada *scene* keempat, citra diambil pada lingkungan luar ruangan. Perekaman *scene* diawali dengan *zoom* minimal kemudian *zoom in* hingga mencapai maksimal dan dilanjutkan dengan *zoom out* hingga kembali pada *zoom* minimal. Ilustrasi pengambilan data dapat dilihat pada 4.6. Data yang didapat untuk *scene* ini sejumlah 650 frame.

### 4.3. Estimasi *focal length* otomatis

Pembahasan pada tahap ini dibagi menjadi dua berdasarkan metode esti-masi *focal length* yang digunakan. Pada masing masing metode, pembahasan diawali dengan jenis data sintetis dan dilanjutkan dengan jenis data citra.



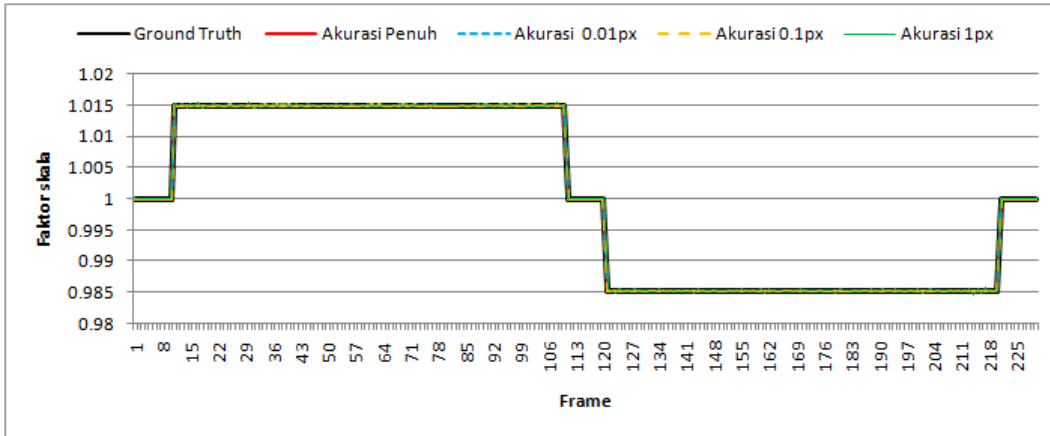
Gambar 4.6: Data *scene* keempat, lingkungan luar ruangan.

#### 4.3.1. Estimasi *focal length* berbasis jarak titik

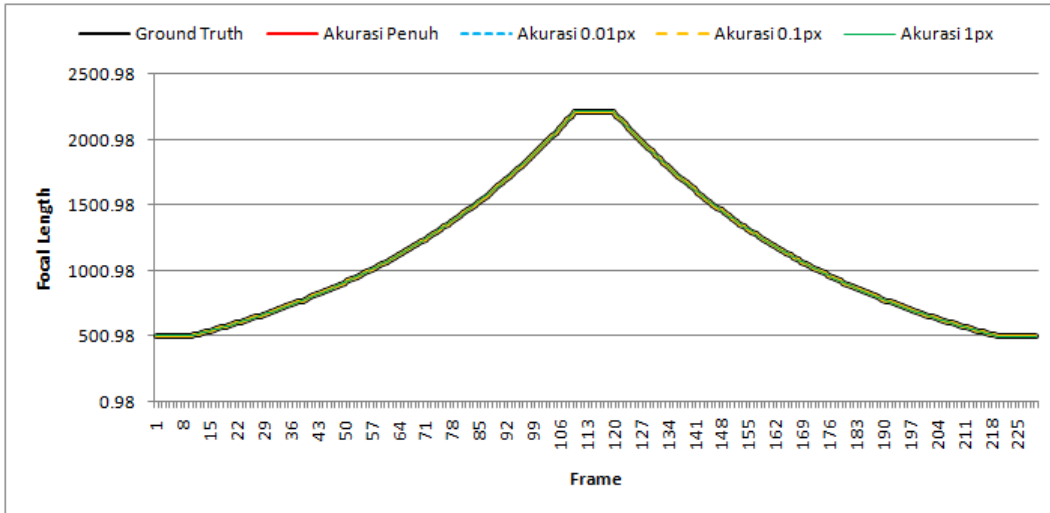
Pada bagian ini estimasi dari data sintetis dilakukan pada empat level akurasi secara terpisah. Korespondensi titik proyeksi dilakukan dengan memasangkan titik-titik dengan label yang sama dari masing-masing hasil proyeksi. Penyajian hasil estimasi dikelompokkan berdasarkan empat skenario pengambilan data sintesis. Hasil estimasi  $s_n$  dan *focal length* masing-masing kamera dari data sintetis yang ditangkap oleh kamera virtual stereo pertama dapat dilihat pada Gambar 4.7 dan Gambar 4.14.

Pada Gambar 4.7, hasil estimasi  $s_n$  hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014995 hingga 1.015003 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.01448 hingga 1.015533. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985218 hingga 0.985226 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.984689 hingga 0.985717. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.7 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.8, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi



Gambar 4.7: Hasil  $s_n$  berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 1.

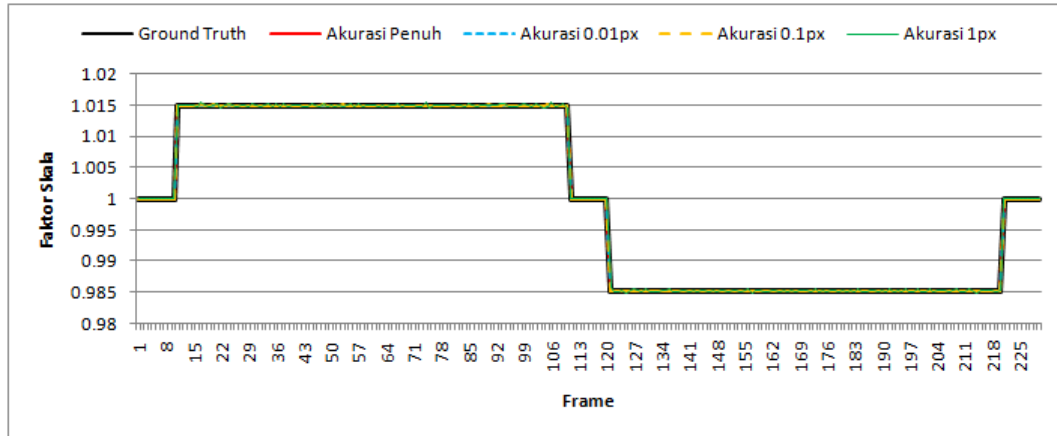


Gambar 4.8: Hasil estimasi *focal length* berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 1.

akurasi penuh 2215.96 pada simulasi akurasi 0.1 piksel dan 2216.033 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 500.0002 dan pada simulasi akurasi 1 piksel nilai hasil estimasi berada pada 499.7288. Gambar 4.8 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

Pada Gambar 4.9, hasil estimasi  $s_n$  hampir identik dengan nilai *ground*

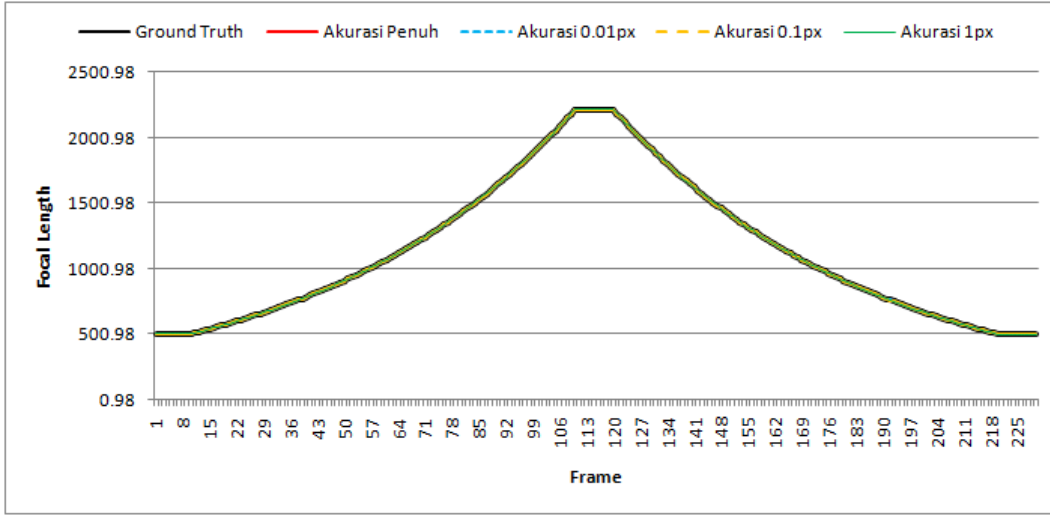




Gambar 4.9: Hasil  $s_n$  berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 2.

*truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014995 hingga 1.015004 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.13778 hingga 1.16679. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985218 hingga 0.985226 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.984756 hingga 0.985715. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.9 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.10, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi akurasi penuh 2215.98624 pada simulasi akurasi 0.1 piksel dan 2216.113 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 500.0007 dan pada

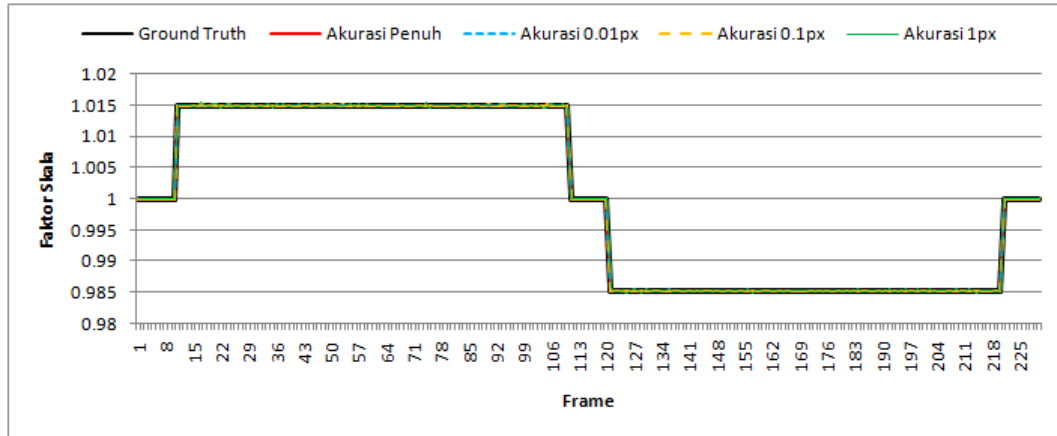


Gambar 4.10: Hasil estimasi *focal length* berbasis jarak titik dari data sintesis menggunakan kamera stereo virtual 2.

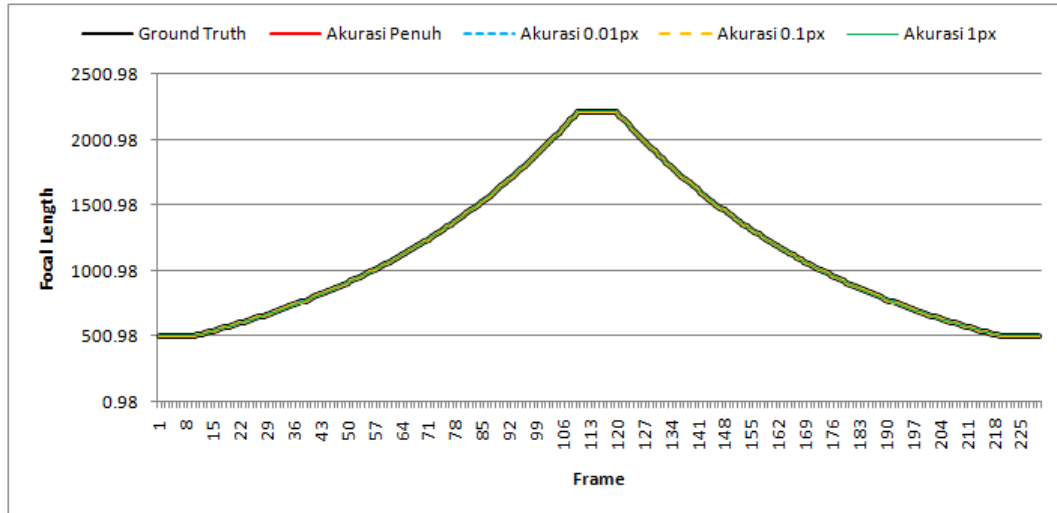
simulasi akurasi 1 piksel nilai hasil estimasi berada pada 499.9492093. Gambar 4.10 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

Pada Gambar 4.11, hasil estimasi  $s_n$  hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014995 hingga 1.015004 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.014507 hingga 1.015523. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985217 hingga 0.985226 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.984777 hingga 0.985703. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.11 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.12, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi



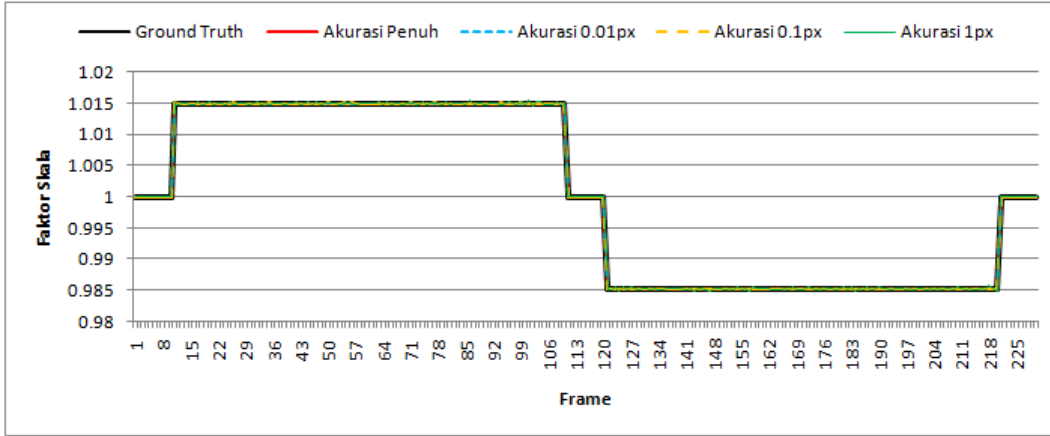
Gambar 4.11: Hasil  $s_n$  berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3.



Gambar 4.12: Hasil estimasi *focal length* berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3.

akurasi penuh 2215.971 pada simulasi akurasi 0.1 piksel dan 2217.189 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 500.003 dan pada simulasi akurasi 1 piksel nilai hasil estimasi berada pada 500.419. Gambar 4.12 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

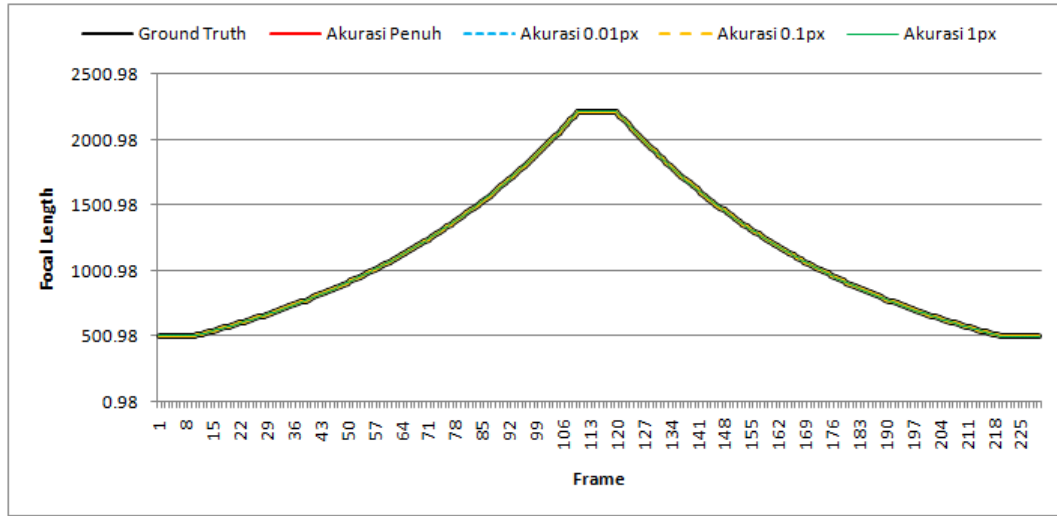
Pada Gambar 4.13, hasil estimasi  $s_n$  hampir identik dengan nilai *ground*



Gambar 4.13: Hasil  $s_n$  berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

*truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014994 hingga 1.015005 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.014639 hingga 1.015421. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985217 hingga 0.985225 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.984731 hingga 0.985697. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.13 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.14, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi akurasi penuh, 2216.033 pada simulasi akurasi 0.1 piksel dan 2216.016 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 499.996 dan

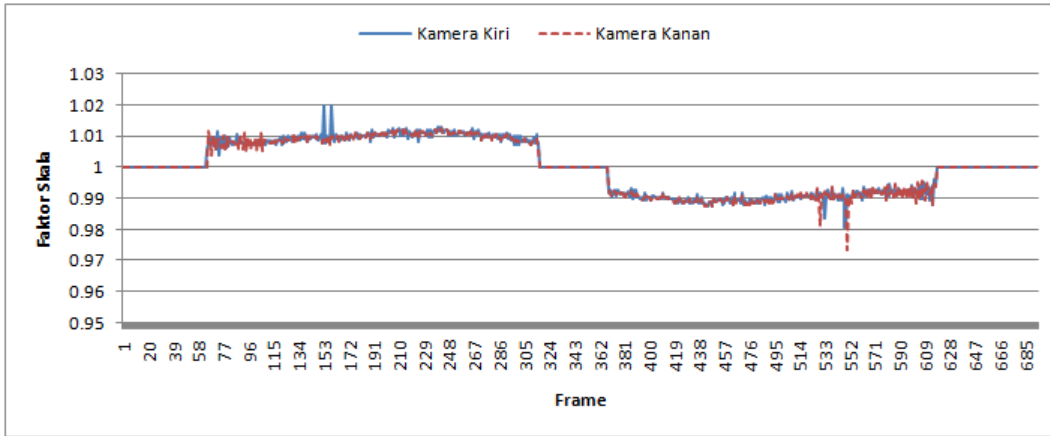


Gambar 4.14: Hasil estimasi *focal length* berbasis jarak titik dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

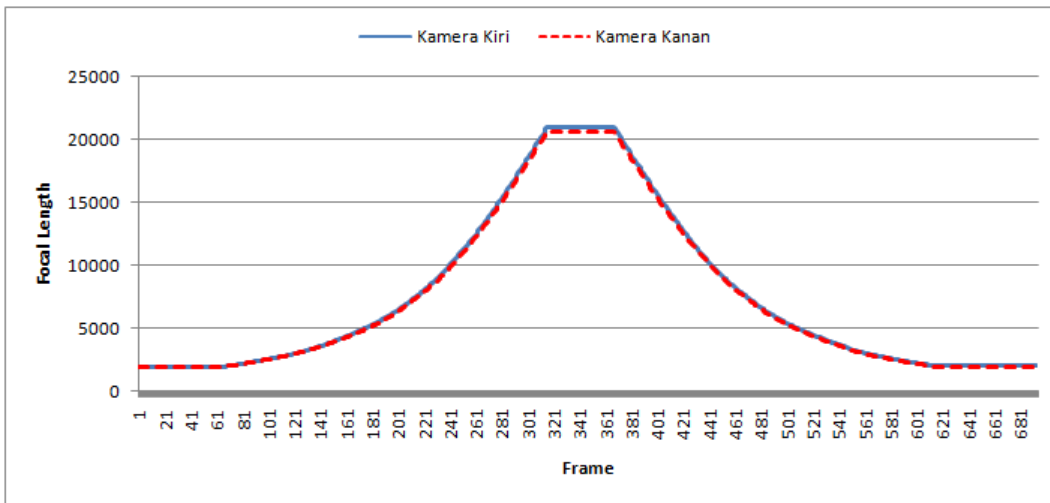
pada simulasi akurasi 1 piksel nilai hasil estimasi berada pada 500.0524 hingga 499.937. Gambar 4.14 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*. Pada jenis data citra estimasi nilai *focal length* dilakukan dengan menggunakan elemen-elemen di dalam fundamental matriks. Pada *scene* pertama, titik-titik *cross corner* digunakan sebagai titik-titik masukan untuk mendapatkan fundamental matriks. Hasil estimasi  $s_n$  dan *focal length* masing-masing kamera untuk *scene* yang mengandung pola papan catur dapat dilihat pada Gambar 4.15 dan Gambar 4.16.

Pada Gambar 4.15 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 65. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 66 hingga 316, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 1.004121 hingga 1.014357. Pada frame 317 hingga 369  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 370 hingga 617 nilai  $s_n$  berada di antara 0.973523 hingga 0.99615181. sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 618 hingga 695 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

Pada Gambar 4.16 nilai *focal length* relatif konstan pada nilai mendekati



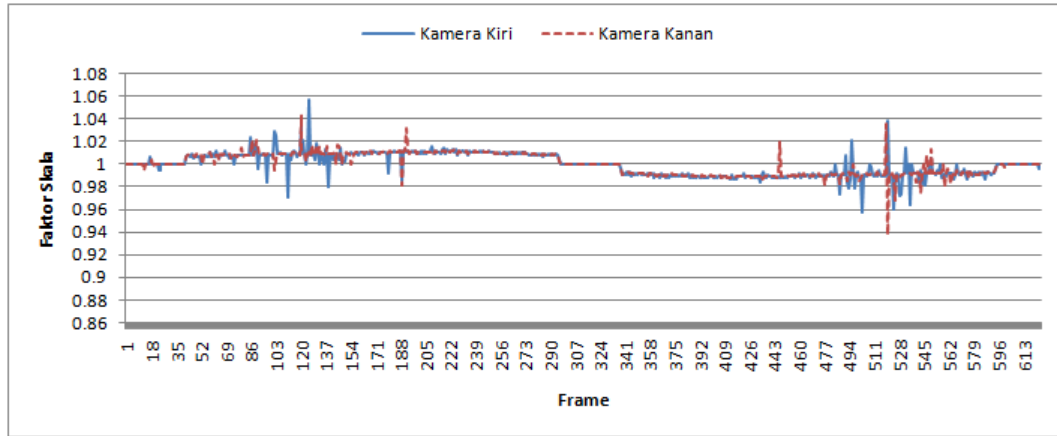
Gambar 4.15: Hasil estimasi  $s_n$  berbasis jarak titik pada *scene* pertama.



Gambar 4.16: Hasil estimasi *focal length* berbasis jarak titik pada *scene* pertama.

1 di frame 1 hingga frame 65. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 66 hingga 316, nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 317 hingga 369. Nilai *focal length* mengalami penurunan pada frame 370 hingga 617 sesuai dengan proses *zoom out*. Pada frame 618 hingga 695 relatif konstan pada 1918.811, sedikit di atas nilai *focal length* pada frame 1 hingga 65.

Pada *scene* kedua hingga keempat, titik yang digunakan adalah titik fitur

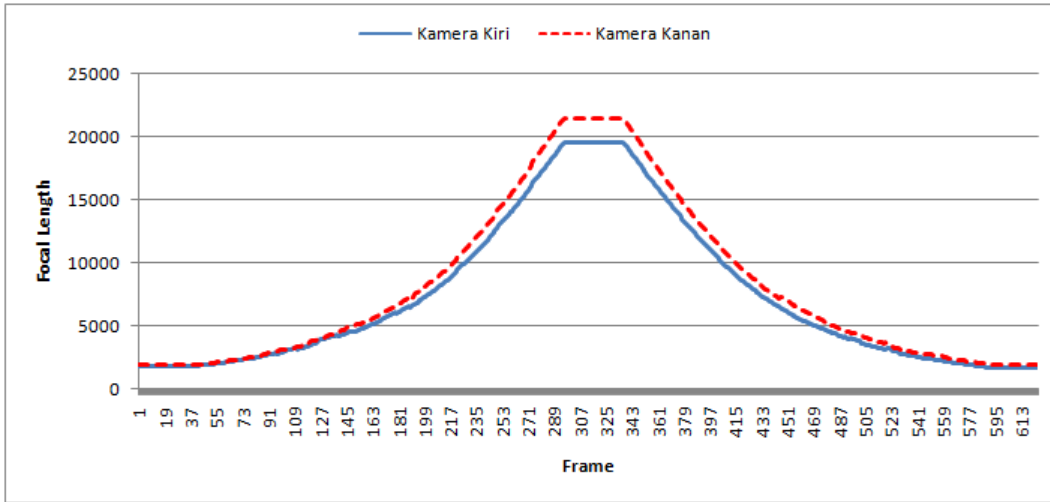


Gambar 4.17: Hasil estimasi  $s_n$  berbasis jarak titik pada *scene* kedua.

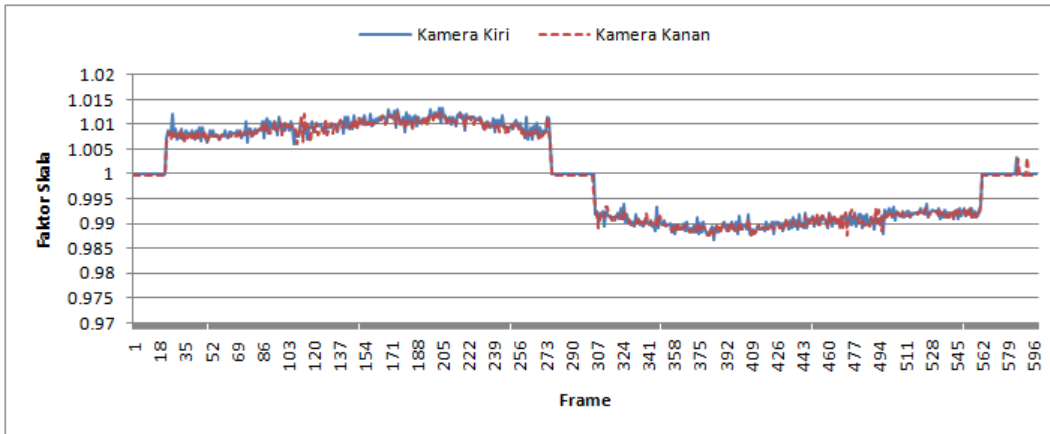
SURF. Penelitian ini menggunakan *Sum of Squared Difference* untuk mendapatkan pasangan titik yang saling berkorespondensi. Hasil estimasi  $s_n$  dan *focal length* masing-masing kamera untuk *scene* kedua dan ketiga dapat dilihat pada Gambar 4.17 dan Gambar 4.22.

Pada Gambar 4.17 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 42. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 43 hingga 289, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 0.998767 hingga 1.016217. Pada frame 290 hingga 337  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 339 hingga 593 nilai  $s_n$  berada di antara 0.984327 hingga 1.00467 sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 594 hingga 625 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

Pada Gambar 4.34 nilai *focal length* relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 42. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 43 hingga 289 nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 290 hingga 337. Nilai *focal length* mengalami penurunan pada frame 338 hingga 593 sesuai dengan proses *zoom out*. Pada frame 594



Gambar 4.18: Hasil estimasi *focal length* berbasis jarak titik pada *scene* kedua.

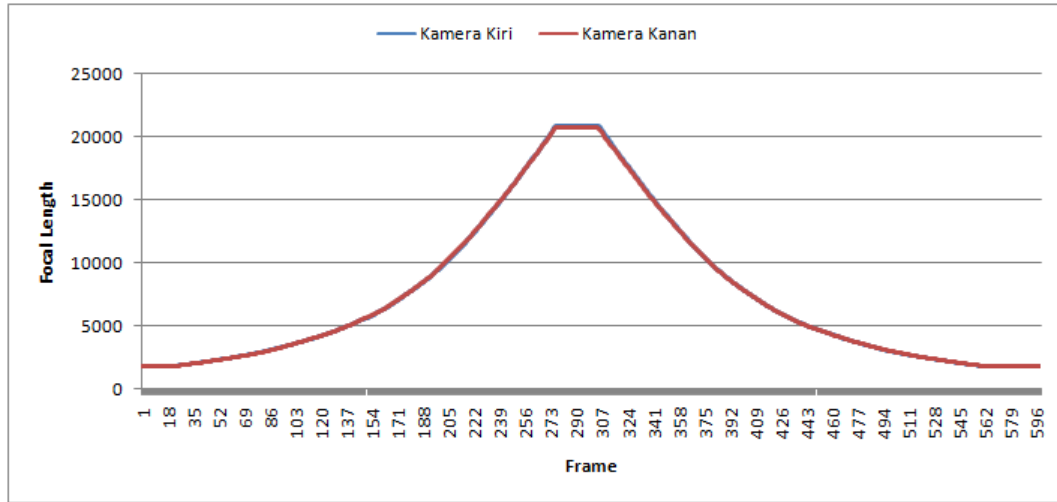


Gambar 4.19: Hasil estimasi  $s_n$  berbasis jarak titik pada *scene* ketiga.

hingga 625 relatif konstan pada 1786.6117, sedikit di bawah nilai *focal length* pada frame 1 hingga 42.

Pada Gambar 4.35 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 24. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 25 hingga 276, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 1.002472 hingga 1.013178. Pada frame 277 hingga 304  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 305 hingga 561 nilai  $s_n$  berada di antara 0.985006 hingga 0.997858 sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 562 hingga 597 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.



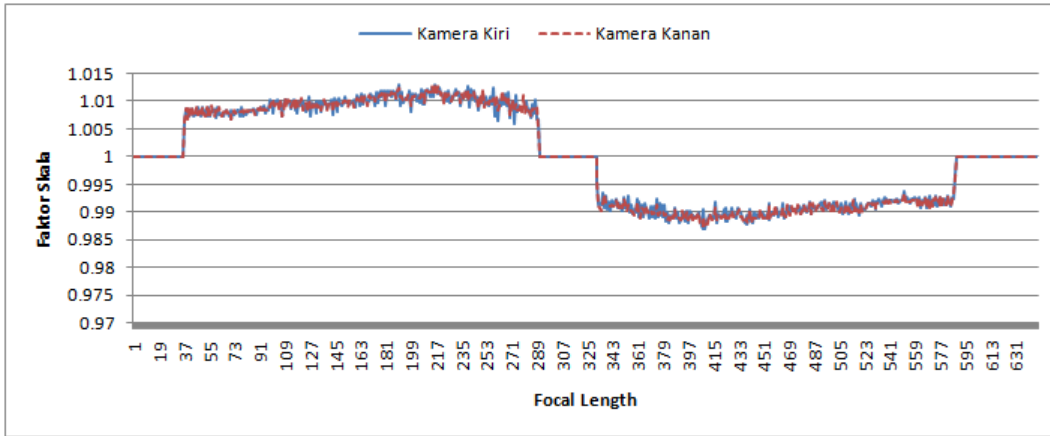


Gambar 4.20: Hasil estimasi *focal length* berbasis jarak titik pada *scene* ketiga.

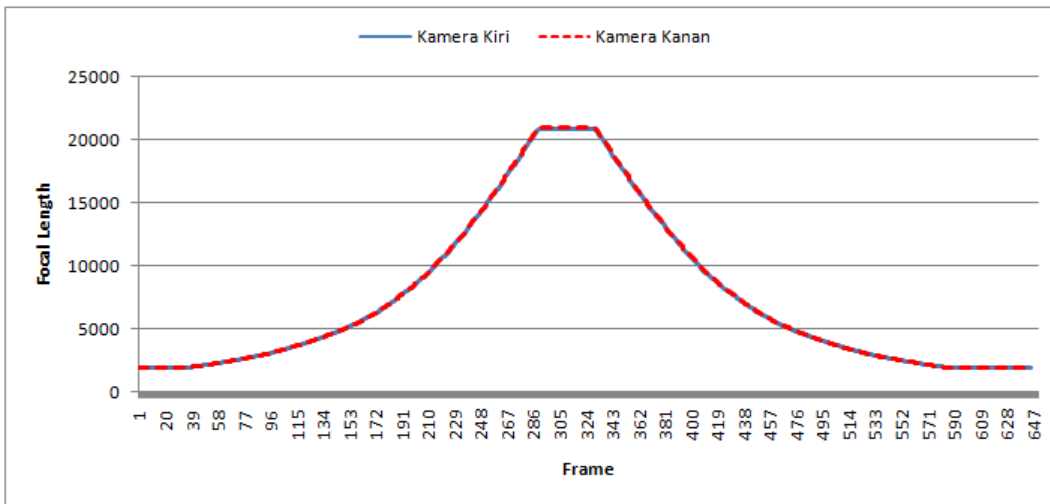
Pada Gambar 4.20 nilai *focal length* relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 24. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 25 hingga 276, nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 277 hingga 304. Nilai *focal length* mengalami penurunan pada frame 305 hingga 561 sesuai dengan proses *zoom out*. Pada frame 562 hingga 597 relatif konstan pada 1889.605, sedikit di atas nilai *focal length* pada frame 1 hingga 24.

Pada Gambar 4.21 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 36. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 37 hingga 294, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 1.00587 hingga 1.013023. Pada frame 295 hingga 336  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 337 hingga 592 nilai  $s_n$  berada di antara 0.986865 hingga 0.999623, sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 593 hingga 650 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

Pada Gambar 4.22 nilai *focal length* relatif konstan pada nilai mendekati



Gambar 4.21: Hasil estimasi  $s_n$  berbasis jarak titik pada *scene* keempat.



Gambar 4.22: Hasil estimasi *focal length* berbasis jarak titik pada *scene* keempat.

1 di frame 1 hingga frame 36. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 37 hingga 294, nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 295 hingga 336. Nilai *focal length* mengalami penurunan pada frame 337 hingga 592 sesuai dengan proses *zoom out*. Pada frame 593 hingga 650 relatif konstan pada 1846.395, sedikit di bawah nilai *focal length* pada frame 1 hingga frame 36.

Dari pengamatan hasil estimasi pada Gambar 4.7 hingga Gambar 4.22, nilai

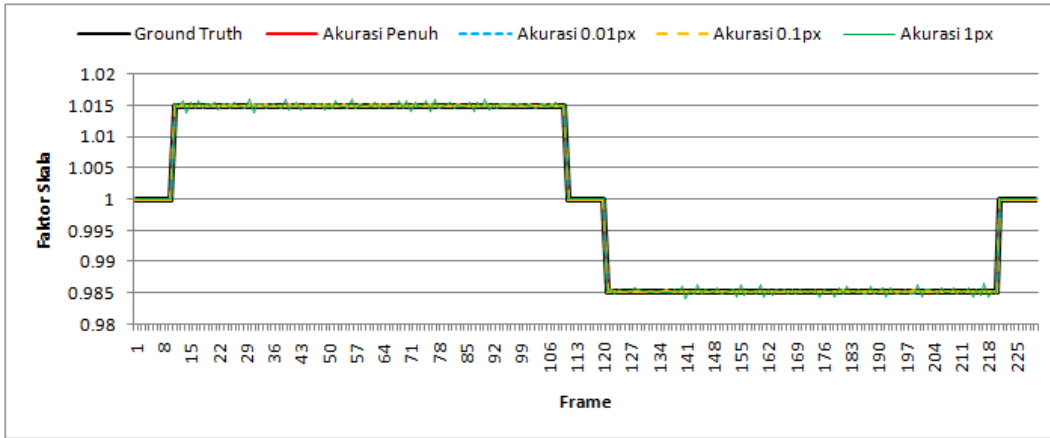
hasil estimasi skala antar-frame dan *focal length* menggunakan fundamental matriks mengikuti terjadi tidaknya proses *zooming*. Untuk skenario proses *zooming* yang serupa, tidak terlihat perbedaan yang signifikan antara berbagai model kamera virtual maupun fisik.

#### 4.3.2. Estimasi *focal length* menggunakan fundamental matriks

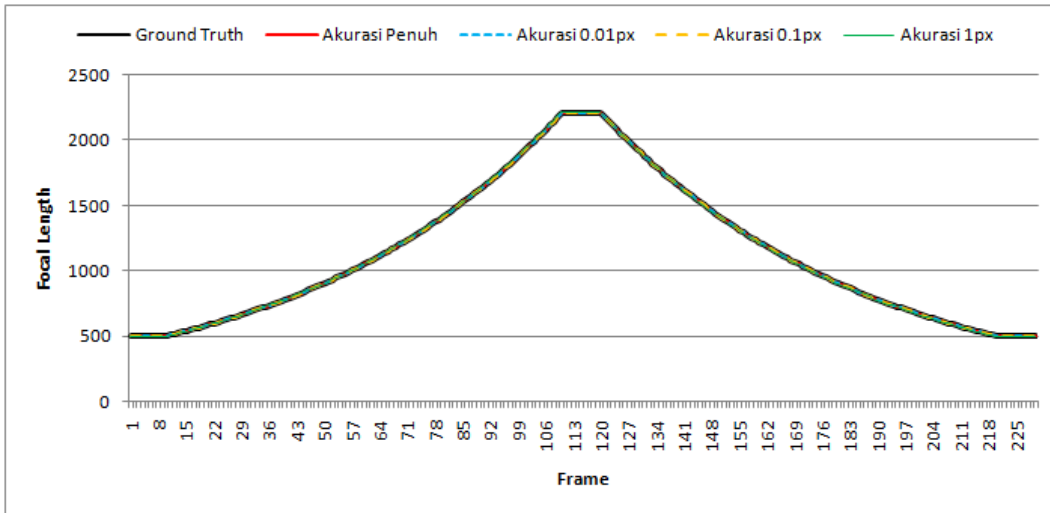
Seperti pada estimasi *focal length* berbasis jarak titik, estimasi *focal length* dari data sintetis dilakukan pada empat level akurasi secara terpisah. Korespondensi titik proyeksi dilakukan dengan memasangkan titik-titik dengan label yang sama dari masing-masing hasil proyeksi. Penyajian hasil estimasi dikelompokkan berdasarkan empat skenario pengambilan data sintesis. Hasil estimasi  $s_n$  dan *focal length* masing-masing kamera dari data sintetis yang ditangkap oleh kamera virtual stereo pertama dapat dilihat pada Gambar 4.23 dan Gambar 4.30.

Pada Gambar 4.23, hasil estimasi  $s_n$  hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014987 hingga 1.015011 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.13778 hingga 1.16101. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985211 hingga 0.985232 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.984097 hingga 0.986392. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.23 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.24, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi



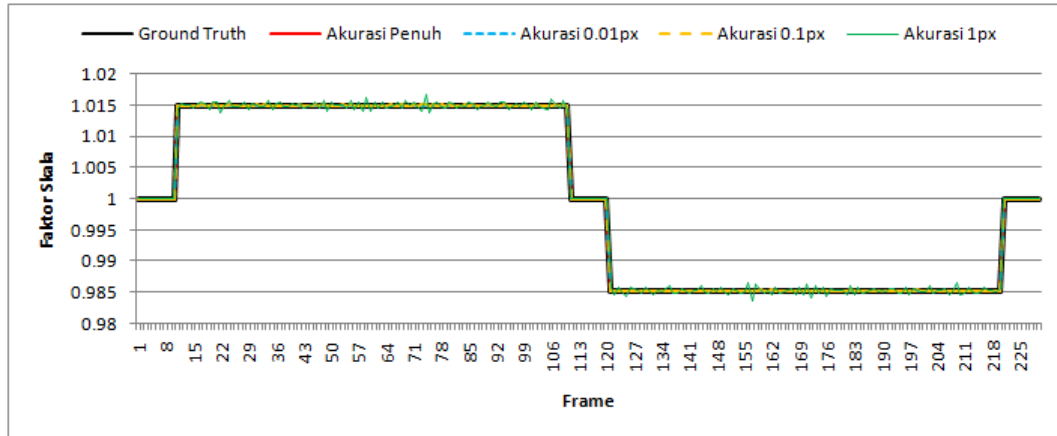
Gambar 4.23: Hasil  $s_n$  menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1.



Gambar 4.24: Hasil estimasi *focal length* menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1.

akurasi penuh 2215.81 pada simulasi akurasi 0.1 piksel dan 2216.608 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 499.9988 dan pada simulasi akurasi 1 piksel nilai hasil estimasi berada pada 499.4401. Gambar 4.24 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

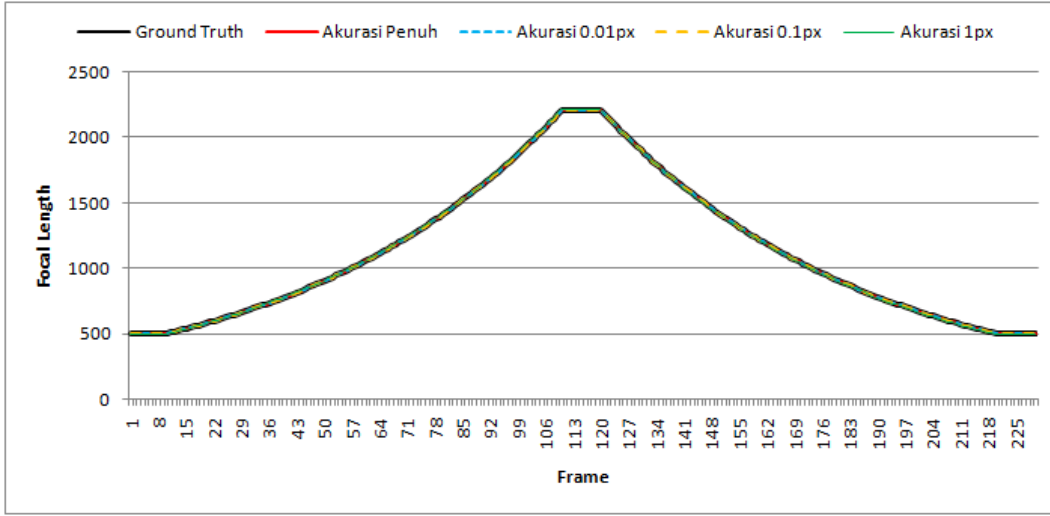
Pada Gambar 4.25, hasil estimasi  $s_n$  hampir identik dengan nilai *ground*



Gambar 4.25: Hasil  $s_n$  menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2.

*truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014987 hingga 1.015014 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.13778 hingga 1.16679. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985209 hingga 0.985234 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.983578 hingga 0.98645. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.25 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.26, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi akurasi penuh 2215.864 pada simulasi akurasi 0.1 piksel dan 2214.058 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 500.0009 dan pada

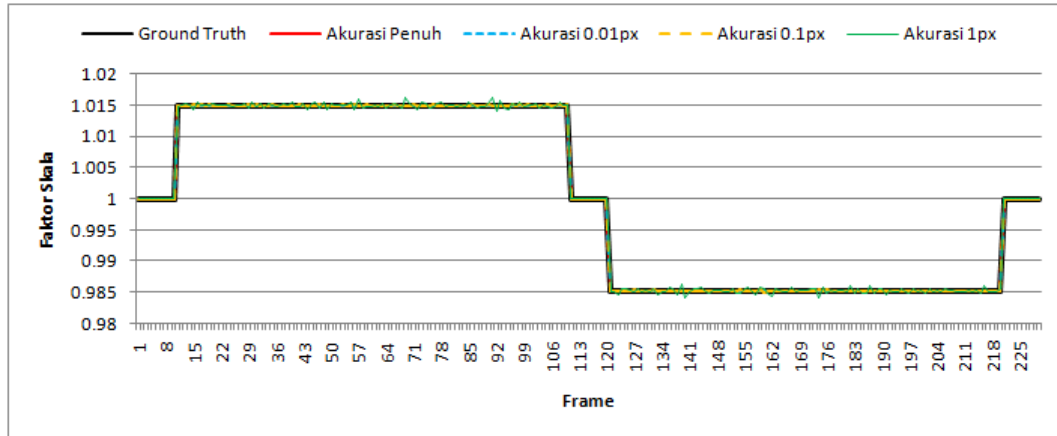


Gambar 4.26: Hasil estimasi *focal length* menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2.

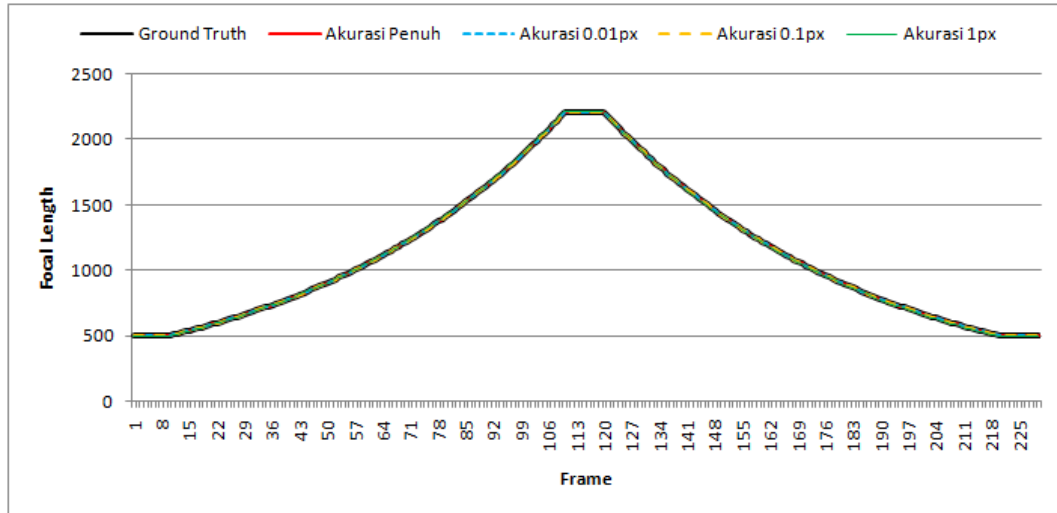
simulasi akurasi 1 piksel nilai hasil estimasi berada pada 499.2448. Gambar 4.26 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

Pada Gambar 4.27, hasil estimasi  $s_n$  hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014987 hingga 1.015011 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.13997 hingga 1.16151. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985211 hingga 0.985236 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.984019 hingga 0.986186. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.27 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.28, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi



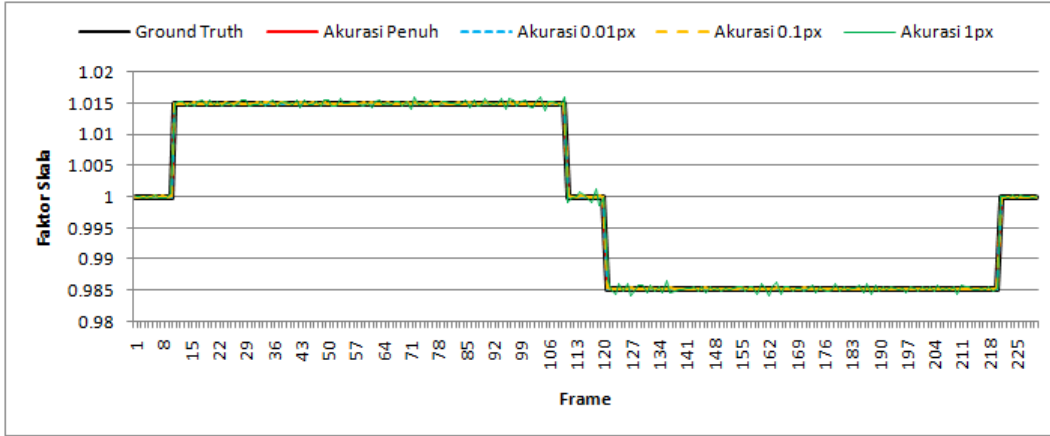
Gambar 4.27: Hasil  $s_n$  menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3.



Gambar 4.28: Hasil estimasi *focal length* menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3.

akurasi penuh 2215.779 pada simulasi akurasi 0.1 piksel dan 2216.69 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil estimasi berada pada 500.0052 dan pada simulasi akurasi 1 piksel nilai hasil estimasi berada pada 499.7571. Gambar 4.28 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

Pada Gambar 4.29, hasil estimasi  $s_n$  hampir identik dengan nilai *ground*

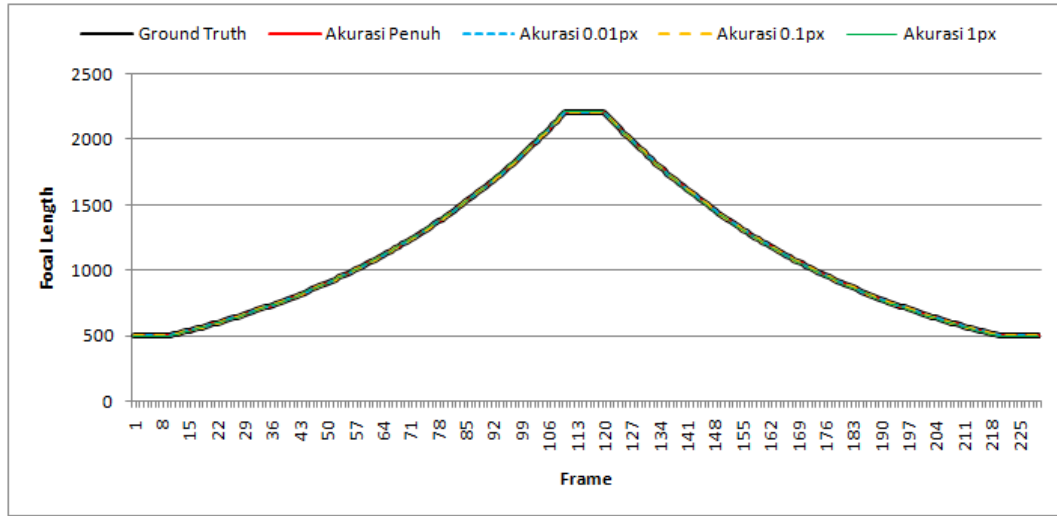


Gambar 4.29: Hasil  $s_n$  menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

*truth*. Pada frame 11 hingga 111 nilai estimasi  $s_n$  mendekati 1.015 dimana simulasi melakukan proses *zoom in*. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 1.014986 hingga 1.015009 dan pada akurasi 1 piksel  $s_n$  berkisar antara 1.13701 hingga 1.16034. Pada frame 121 hingga 220 nilai estimasi  $s_n$  berada sekitar 0.98522 mengikuti proses *zoom out* dari simulasi. Pada simulasi akurasi data 0.01 piksel, nilai estimasi  $s_n$  berkisar antara 0.985207 hingga 0.985235 dan pada akurasi 1 piksel  $s_n$  berkisar antara 0.98398 hingga 0.986511. Pada saat *ground truth* bernilai 1 di awal, pertengahan dan akhir *sequence*, nilai hasil estimasi juga bernilai 1 yang berarti tidak terjadi proses *zooming*. Gambar 4.29 menunjukkan rentang variasi estimasi  $s_n$  semakin membesar ketika akurasi korespondensi menurun.

Pada Gambar 4.30, hasil estimasi *focal length* hampir identik dengan nilai *ground truth*. Pada frame 11 hingga 111 nilai estimasi *focal length* mengalami peningkatan dimana simulasi melakukan proses *zoom in*. Peningkatan *focal length* berhenti pada frame 112 hingga 120 dengan nilai 2216.023 pada simulasi akurasi penuh 2215.992 hingga 2216.031 pada simulasi akurasi 0.1 piksel dan 2215.458 hingga 2219.312 pada akurasi piksel. Pada frame 121 hingga 220 nilai estimasi *focal length* mengalami penurunan mengikuti proses *zoom out* dari simulasi. Pada saat *ground truth* bernilai 500 di akhir *sequence*, nilai hasil estimasi mendekati nilai 500. Pada simulasi akurasi 0.01 piksel, nilai hasil



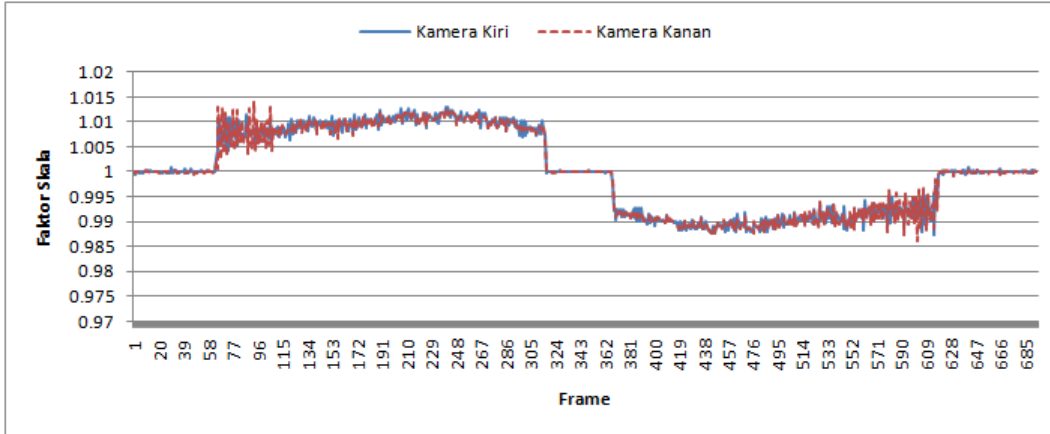


Gambar 4.30: Hasil estimasi *focal length* menggunakan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

estimasi berada pada 499.9928 hingga 499.9992 dan pada simulasi akurasi 1 piksel nilai hasil estimasi berada pada 500.0524 hingga 500.2644. Gambar 4.30 menunjukkan bahwa semakin menurun tingkat akurasi, nilai *focal length* hasil estimasi semakin menjauh dari *ground truth*.

Pada jenis data citra estimasi nilai *focal length* dilakukan dengan menggunakan elemen-elemen di dalam fundamental matriks. Pada *scene* pertama, titik-titik *cross corner* digunakan sebagai titik-titik masukan untuk mendapatkan fundamental matriks. Hasil estimasi  $s_n$  dan *focal length* masing-masing kamera untuk *scene* yang mengandung pola papan catur dapat dilihat pada Gambar 4.31 dan Gambar 4.32.

Pada Gambar 4.31 nilai  $s_n$  relatif konstan pada nilai 1 di frame 1 hingga frame 65. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 66 hingga 316, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 1.003845 hingga 1.019698. Pada frame 317 hingga 369  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 370 hingga 617 nilai  $s_n$  berada di antara 0.987280 hingga 0.998485. sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 618 hingga 695 relatif konstan



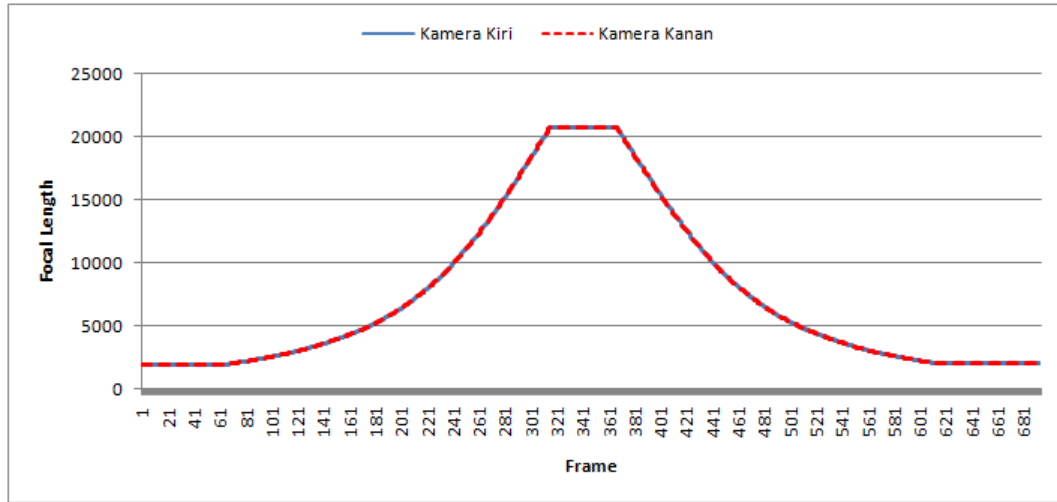
Gambar 4.31: Hasil estimasi  $s_n$  menggunakan fundamental matriks pada *scene* pertama.

pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

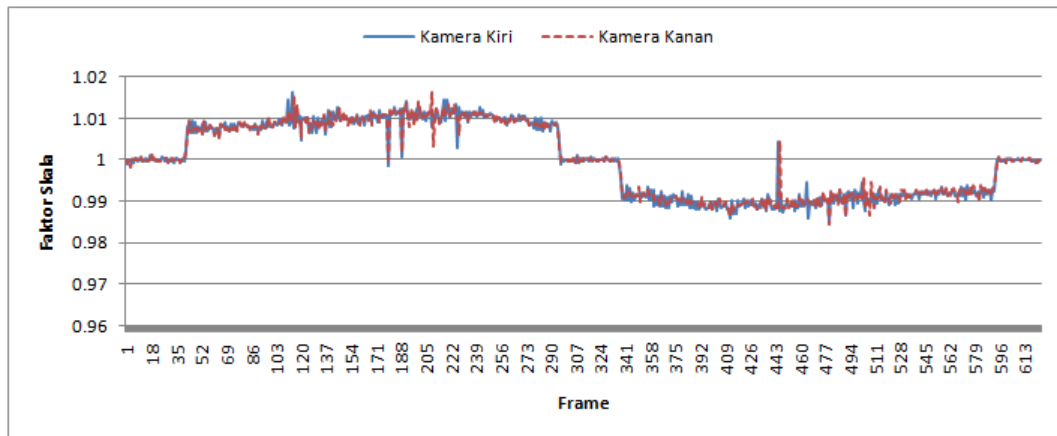
Pada Gambar 4.32 nilai *focal length* relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 65. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 66 hingga 316, nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 317 hingga 369. Nilai *focal length* mengalami penurunan pada frame 370 hingga 617 sesuai dengan proses *zoom out*. Pada frame 618 hingga 695 relatif konstan pada 1928.83, sedikit di atas nilai *focal length* pada frame 1 hingga 65.

Pada *scene* kedua hingga keempat, titik yang digunakan adalah titik fitur SURF. Penelitian ini menggunakan *Sum of Squared Difference* untuk mendapatkan pasangan titik yang saling berkorespondensi. Hasil estimasi  $s_n$  dan *focal length* masing-masing kamera untuk *scene* kedua hingga keempat dapat dilihat pada Gambar 4.33 dan Gambar 4.38.

Pada Gambar 4.33 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 42. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 43 hingga 289, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 0.998767 hingga 1.016217. Pada frame



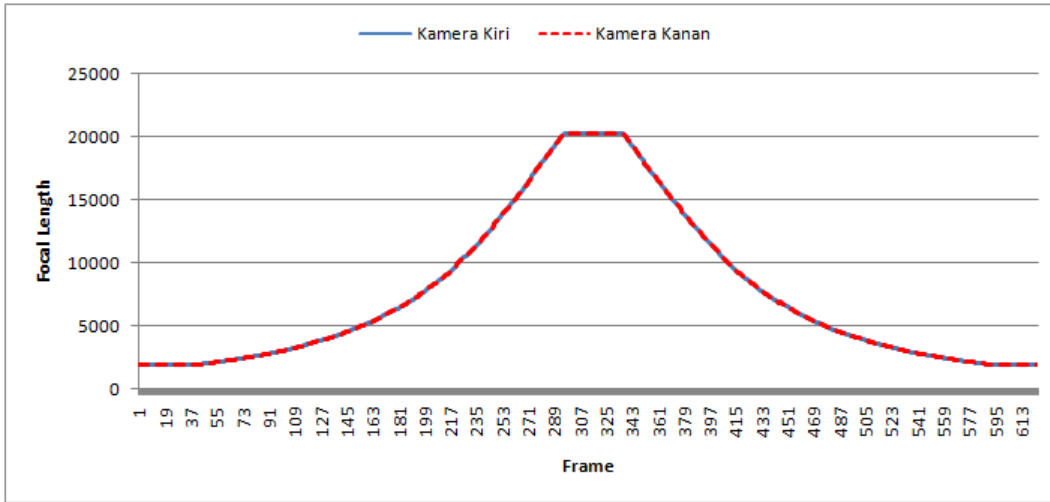
Gambar 4.32: Hasil estimasi *focal length* menggunakan fundamental matriks pada *scene* pertama.



Gambar 4.33: Hasil estimasi  $s_n$  menggunakan fundamental matriks pada *scene* kedua.

290 hingga 337  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 339 hingga 593 nilai  $s_n$  berada di antara 0.984327 hingga 1.00467 sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 594 hingga 625 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

Pada Gambar 4.34 nilai *focal length* relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 42. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 43 hingga 289

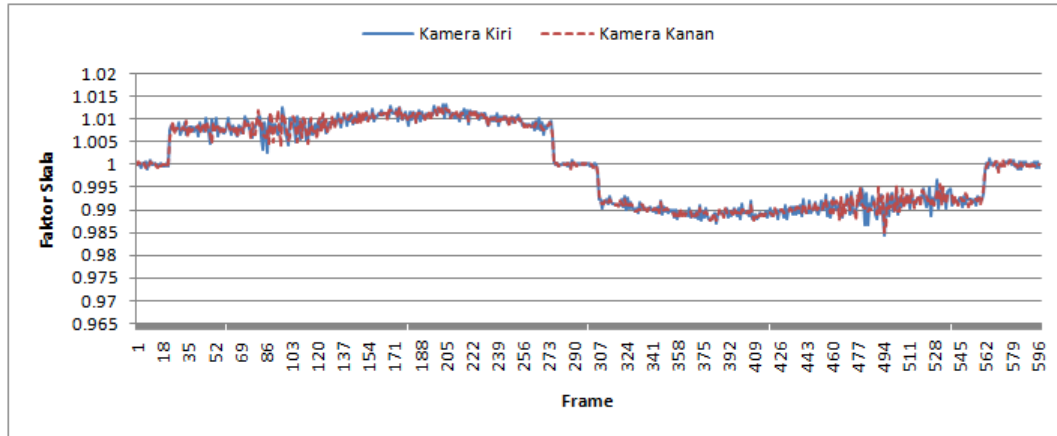


Gambar 4.34: Hasil estimasi *focal length* menggunakan fundamental matriks pada *scene* kedua.

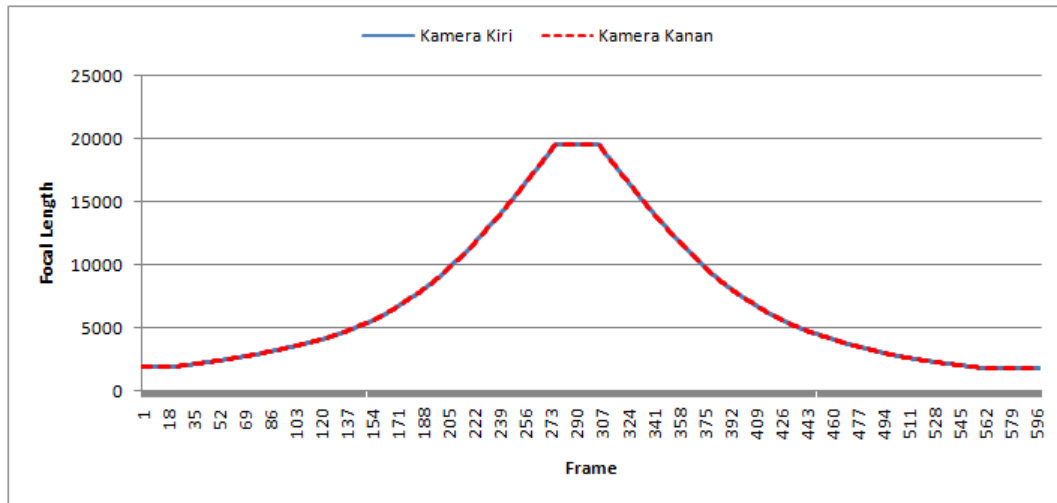
nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 290 hingga 337. Nilai *focal length* mengalami penurunan pada frame 338 hingga 593 sesuai dengan proses *zoom out*. Pada frame 594 hingga 625 relatif konstan pada 1845.7528, sedikit di bawah nilai *focal length* pada frame 1 hingga 42.

Pada Gambar 4.35 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 24. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 25 hingga 276, citra mengalami proses *zoom in* dan mengakibatkan nilai  $s_n$  antara 1.002472 hingga 1.013178. Pada frame 277 hingga 304  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 305 hingga 561 nilai  $s_n$  berada di antara 0.985006 hingga 0.997858 sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 562 hingga 597 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

Pada Gambar 4.36 nilai *focal length* relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 24. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 25 hingga 276,



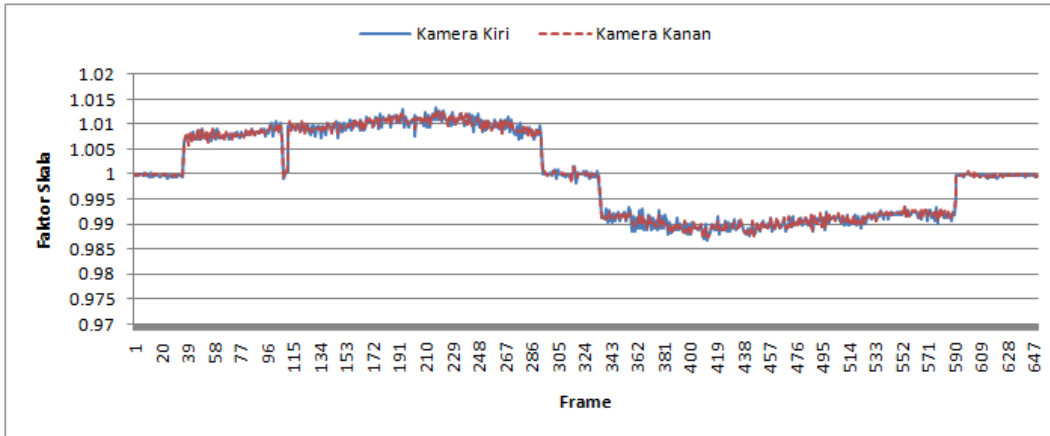
Gambar 4.35: Hasil estimasi  $s_n$  menggunakan fundamental matriks pada *scene* ketiga.



Gambar 4.36: Hasil estimasi *focal length* menggunakan fundamental matriks pada *scene* ketiga.

nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 277 hingga 304. Nilai *focal length* mengalami penurunan pada frame 305 hingga 561 sesuai dengan proses *zoom out*. Pada frame 562 hingga 597 relatif konstan pada 1709.103, sedikit di bawah nilai *focal length* pada frame 1 hingga 24.

Pada Gambar 4.37 nilai  $s_n$  relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 36. Hal ini sesuai dengan kondisi awal rekaman ketika tidak terjadi *zooming*. Pada frame 37 hingga 294, citra mengalami proses *zoom in*

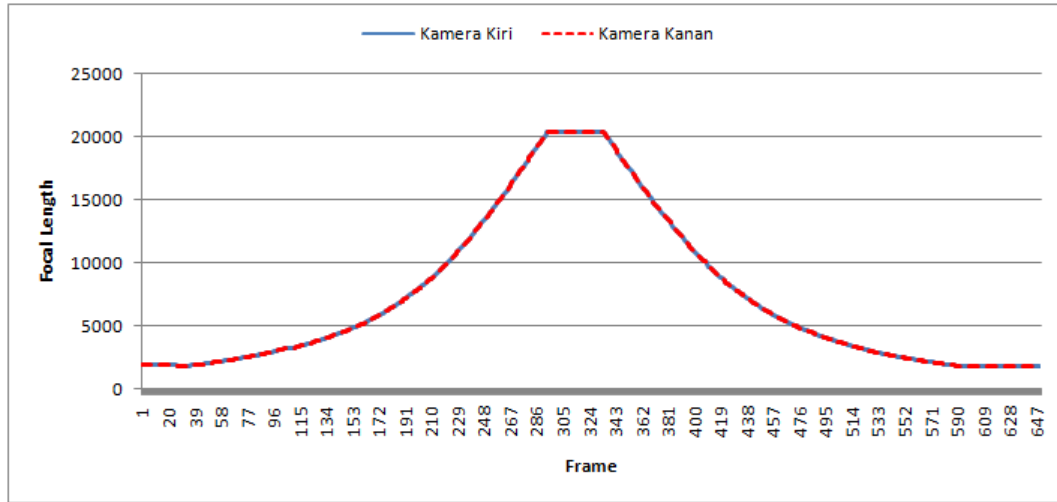


Gambar 4.37: Hasil estimasi  $s_n$  menggunakan fundamental matriks pada *scene* keempat.

dan mengakibatkan nilai  $s_n$  antara 0.999172 hingga 1.013424. Pada frame 295 hingga 336  $s_n$  relatif konstan mendekati 1 sesuai dengan kondisi berhenti sejenak setelah mencapai *zoom* maksimal. Pada frame 337 hingga 592 nilai  $s_n$  berada di antara 0.986671 hingga 0.999623. sesuai dengan proses *zoom out* yang direkam oleh kamera. Nilai  $s_n$  pada frame 593 hingga 650 relatif konstan pada nilai 1 sesuai dengan berhenti sejenak setelah mencapai *zoom* minimal.

Pada Gambar 4.38 nilai *focal length* relatif konstan pada nilai mendekati 1 di frame 1 hingga frame 36. Hal ini menunjukkan pada frame-frame tersebut nilai *focal length* mendekati pada nilai yang disajikan pada Tabel 4.4 dan hampir tidak mengalami perubahan *focal length*. Pada frame 37 hingga 294, nilai *focal length* mengalami peningkatan yang disebabkan oleh proses *zoom in*. Nilai *focal length* mencapai maksimal dan tidak mengalami perubahan yang signifikan pada frame 295 hingga 336. Nilai *focal length* mengalami penurunan pada frame 337 hingga 592 sesuai dengan proses *zoom out*. Pada frame 593 hingga 650 relatif konstan pada 1801.775, sedikit di bawah nilai *focal length* pada frame 1 hingga frame 36.

Dari pengamatan hasil estimasi pada Gambar 4.23 hingga Gambar 4.38, nilai hasil estimasi skala antar-frame dan *focal length* menggunakan fundamental matriks mengikuti terjadi tidaknya proses *zooming*. Untuk skenario proses *zooming* yang serupa, tidak terlihat perbedaan yang signifikan antara berbagai



Gambar 4.38: Hasil estimasi *focal length* menggunakan fundamental matriks pada *scene* keempat.

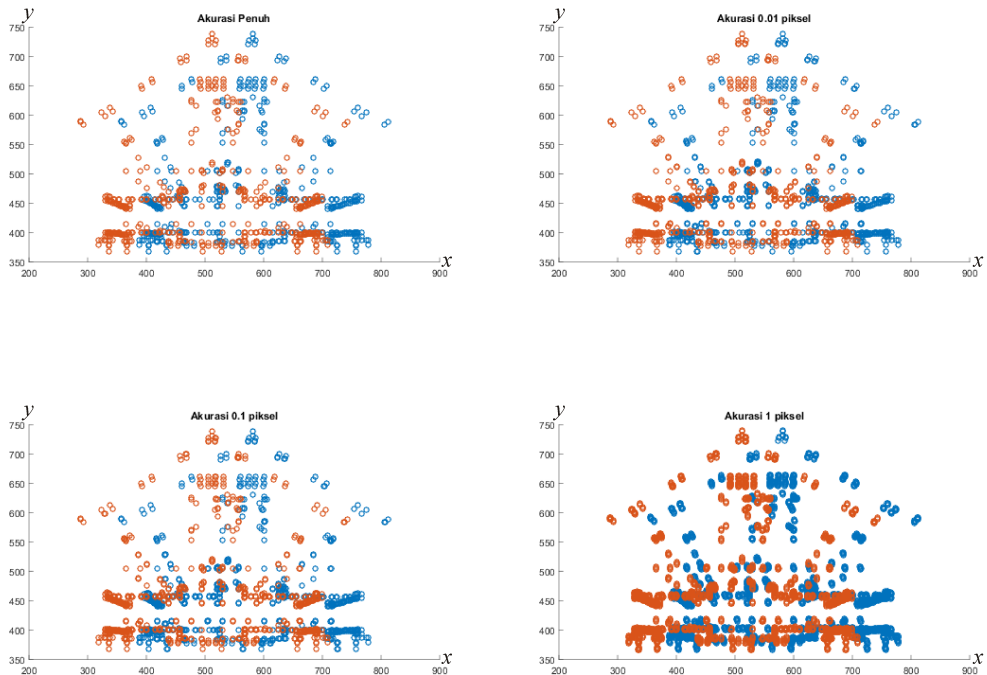
model kamera virtual maupun fisik.

#### 4.4. Normalisasi koordinat titik

Pembahasan pada subbab ini dibagi menjadi dua berdasarkan metode estimasi *center of expansion* yang digunakan. Pada masing-masing metode, pembahasan diawali dengan jenis data sintetis dan dilanjutkan dengan jenis data citra.

##### 4.4.1. Normalisasi berbasis garis ekspansi

Pada metode ini faktor skala antar frame  $s_n$  yang digunakan adalah nilai faktor skala yang didapat menggunakan metode berbasis jarak antar titik. Untuk jenis data sintetis, normalisasi data hasil proyeksi pada setiap kamera dilakukan secara terpisah dengan nilai *focal length* dan lokasi *center of expansion* hasil estimasi masing-masing. Normalisasi pada masing-masing simulasi akurasi proyeksi juga dilakukan secara terpisah. Penyajian data dikelompokkan berdasarkan jenis kamera dan skenario *zooming*. Bagian ini menyajikan hasil normalisasi dari titik-titik hasil proyeksi kamera pertama dan kedua pada masing-masing kamera stereo virtual di dalam satu bidang. Titik-titik berwarna biru merupakan hasil normalisasi titik-titik yang terproyeksi pada

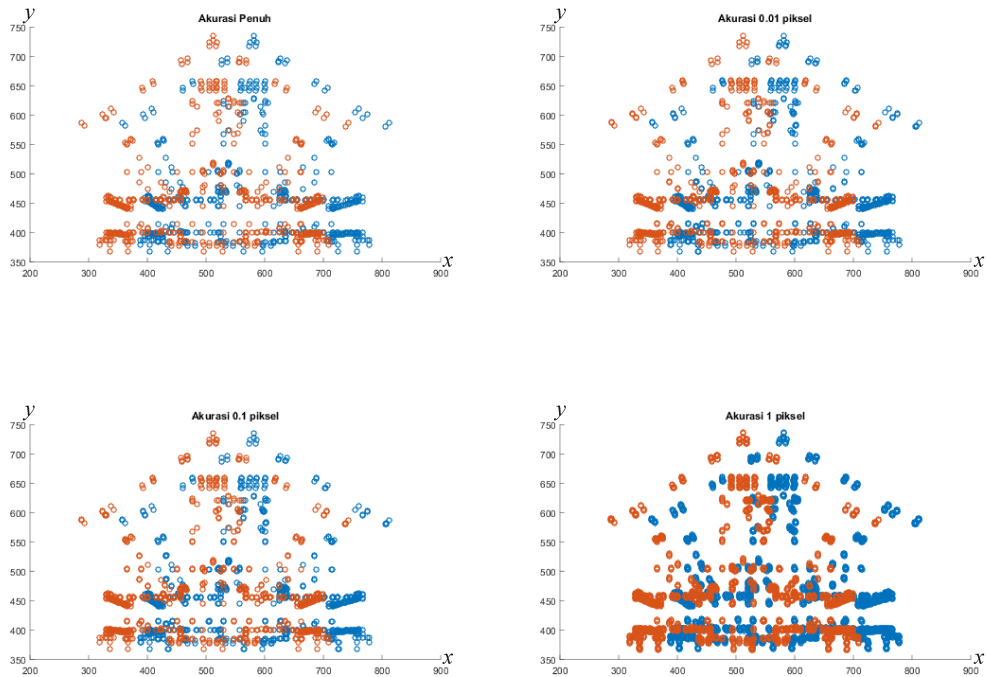


Gambar 4.39: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 1.

kamera pertama sedangkan titik-titik merah merupakan hasil normalisasi titik-titik yang terproyeksi pada kamera kedua. Hasil normalisasi disajikan pada Gambar 4.39 hingga Gambar 4.42.

Gambar 4.39 menunjukkan bahwa pada simulasi akurasi penuh hingga 0.1 piksel, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama berhasil ternormalisasi pada posisi yang relatif sama, namun seiring dengan menurunnya akurasi, persebaran hasil normalisasi dari titik-titik yang sama ikut membesar. Sedangkan untuk akurasi 1 piksel, dapat diamati bahwa titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Hal ini berarti bahwa untuk proyeksi pada kamera stereo virtual dengan *focal length* dan posisi *image center* yang sama, kesalahan korespondensi yang signifikan mulai terjadi pada tingkat akurasi 1 piksel.

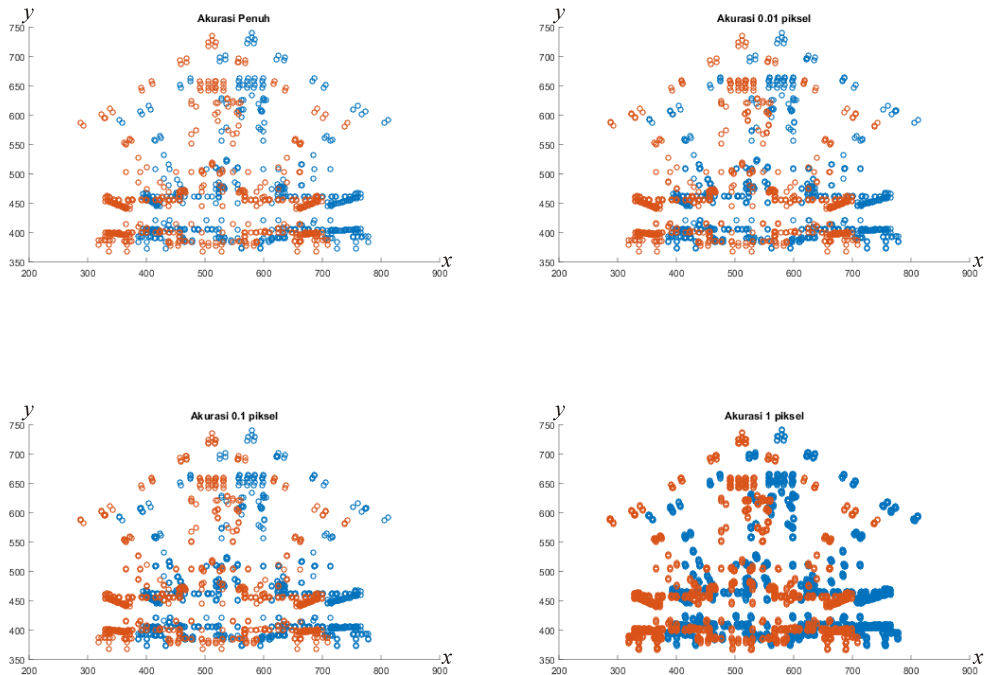




Gambar 4.40: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 2.

Gambar 4.40 menunjukkan bahwa pada simulasi akurasi penuh hingga 0.1 piksel, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama berhasil ternormalisasi pada posisi yang relatif sama namun seiring dengan menurunnya akurasi, persebaran hasil normalisasi dari titik-titik yang sama ikut membesar. Sedangkan untuk akurasi 1 piksel, dapat diamati bahwa titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Hal ini berarti bahwa untuk proyeksi pada kamera stereo dengan posisi *image center* yang sama namun terdapat perbedaan  $f_x$  dan  $f_y$  pada masing-masing kamera, kesalahan korespondensi yang signifikan mulai terjadi pada tingkat akurasi 1 piksel.

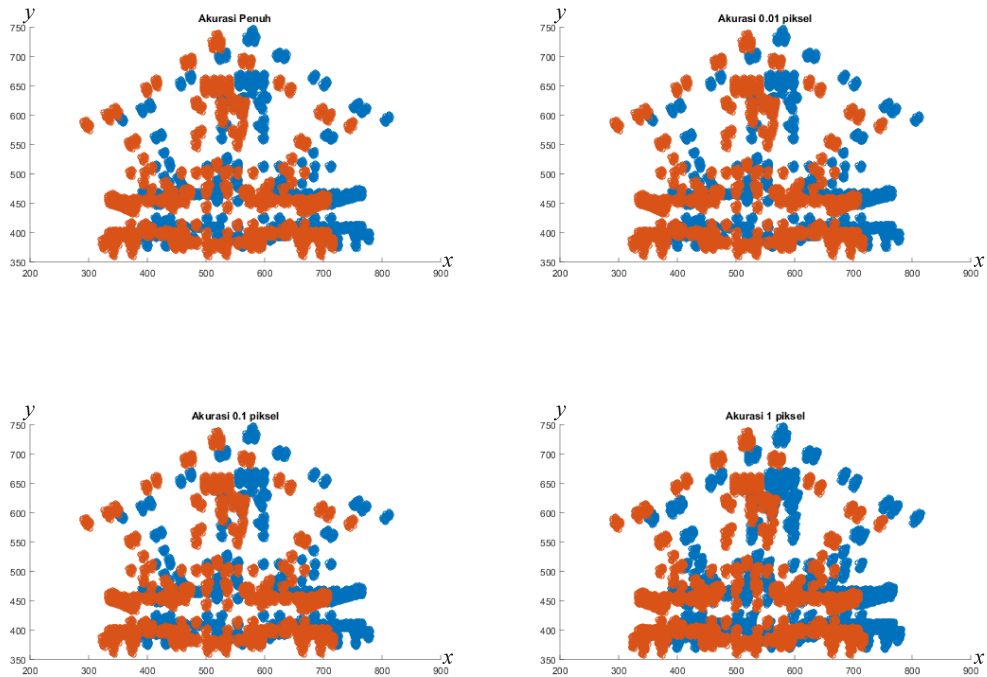
Gambar 4.41 menunjukkan bahwa pada simulasi akurasi penuh hingga 0.1 piksel, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama berhasil ternormalisasi pada posisi yang relatif



Gambar 4.41: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3.

sama namun seiring dengan menurunnya akurasi, persebaran hasil normalisasi dari titik-titik yang sama ikut membesar. Sedangkan untuk akurasi 1 piksel, dapat diamati bahwa titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Hal ini berarti bahwa untuk proyeksi pada kamera stereo di mana masing-masing kamera memiliki parameter intrinsik yang berbeda, kesalahan korespondensi yang signifikan mulai terjadi pada tingkat akurasi 1 piksel.

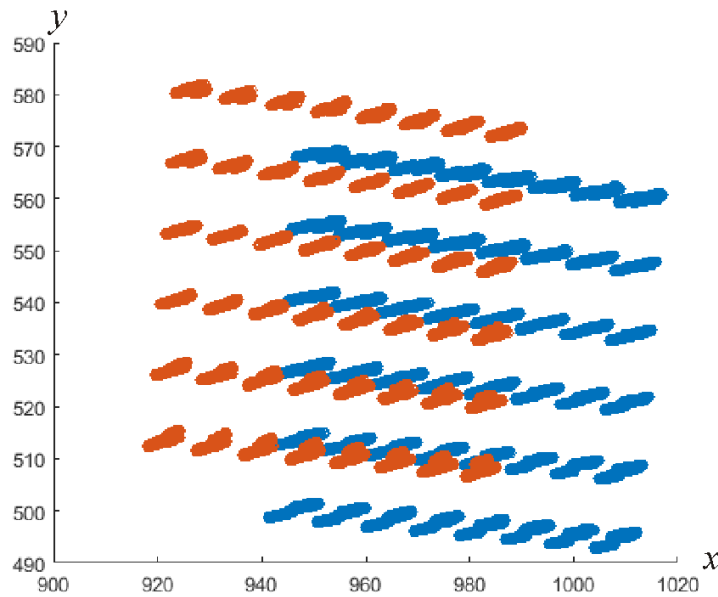
Gambar 4.42 menunjukkan bahwa pada semua level simulasi akurasi, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Jika dibandingkan dengan hasil-hasil pada normalisasi sebelumnya, perubahan posisi *image center* sangat menentukan akurasi dari korespondensi titik-titik dua dimensi hasil proyeksi.



Gambar 4.42: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

Pada jenis data citra, bidang target normalisasi adalah frame pertama di mana level *zoom* pada frame pertama merupakan *zoom minimal*. Normalisasi dilakukan menggunakan frame dengan *zoom* minimal karena pada level *zoom* tersebut parameter intrinsik dan ekstrinsik kamera diketahui dari hasil kalibrasi awal. Bagian menyajikan hasil normalisasi dari titik-titik hasil proyeksi kamera kiri dan kanan pada masing-masing kamera stereo di dalam satu bidang. Titik-titik berwarna biru merupakan hasil normalisasi titik-titik objek atau lingkungan yang terproyeksi pada kamera kiri sedangkan titik-titik merah merupakan hasil normalisasi titik-titik yang terproyeksi pada kamera kanan.

Gambar 4.43 menunjukkan bahwa data hasil normalisasi dari objek yang sama tersebar secara vertikal horizontal dengan persebaran yang cukup luas. Dilhat dari bentuk persebarannya, diperkirakan titik *center of expansion* hasil estimasi terpaut lebih jauh secara horizontal daripada secara vertikal. Hal ini



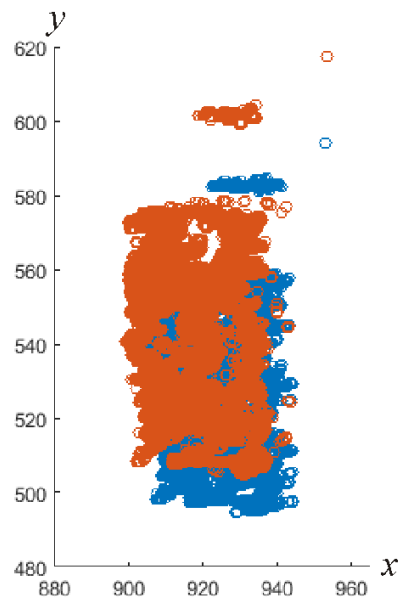
Gambar 4.43: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* pertama.

berpotensi mengakibatkan error pada hasil rekonstruksi yang cukup besar.

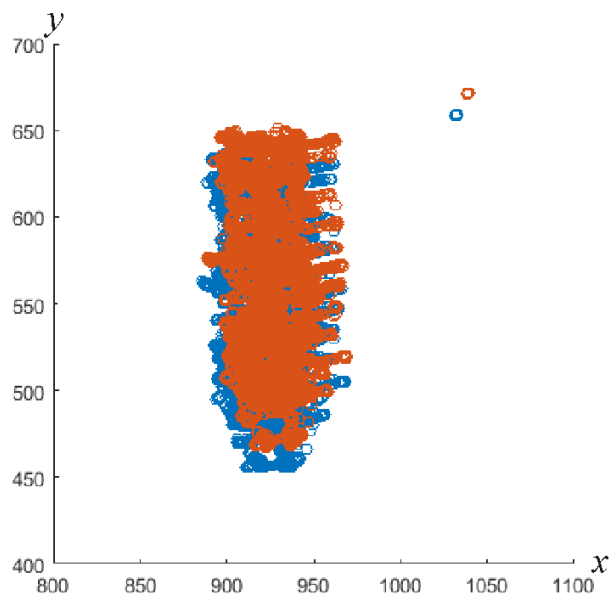
Gambar 4.44 kumpulan titik hasil normalisasi memberikan bentuk yang menyerupai persegi panjang. Tampak bahwa data hasil normalisasi dari objek yang sama tersebar secara vertikal horizontal dengan persebaran yang cukup luas. Dilhat dari bentuk persebarannya, diperkirakan titik *center of expansion* hasil estimasi terpaut lebih jauh secara horizontal daripada secara vertikal. Hal ini berpotensi mengakibatkan error pada hasil rekonstruksi yang cukup besar.

Gambar 4.45 kumpulan titik hasil normalisasi memberikan bentuk yang menyerupai botol terbalik. Botol terbalik karena sumbu  $y$  positif pada citra berlawanan arah dengan sumbu  $y$  positif pada grafik. Tampak bahwa data hasil normalisasi dari objek yang sama tersebar secara vertikal horizontal dengan persebaran yang cukup luas. Dilhat dari bentuk persebarannya, diperkirakan titik *center of expansion* hasil estimasi terpaut lebih jauh secara horizontal daripada secara vertikal. Hal ini berpotensi mengakibatkan error pada hasil rekonstruksi yang cukup besar.

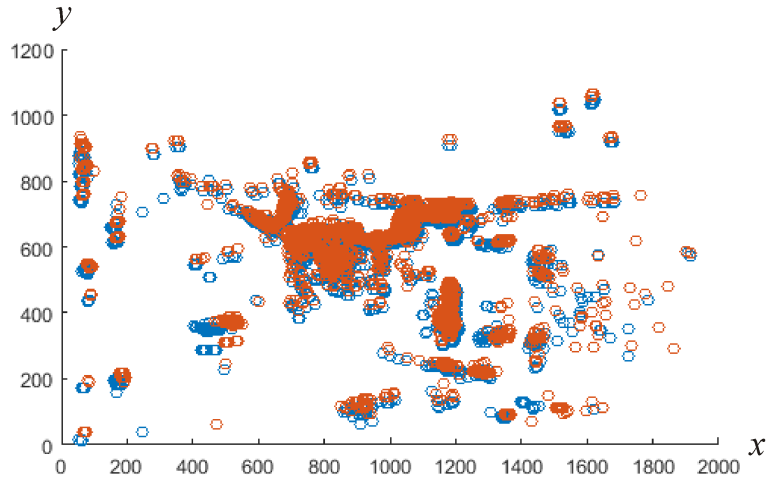
Pada Gambar 4.46 tampak bahwa data hasil normalisasi dari objek yang



Gambar 4.44: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* kedua.



Gambar 4.45: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* ketiga.



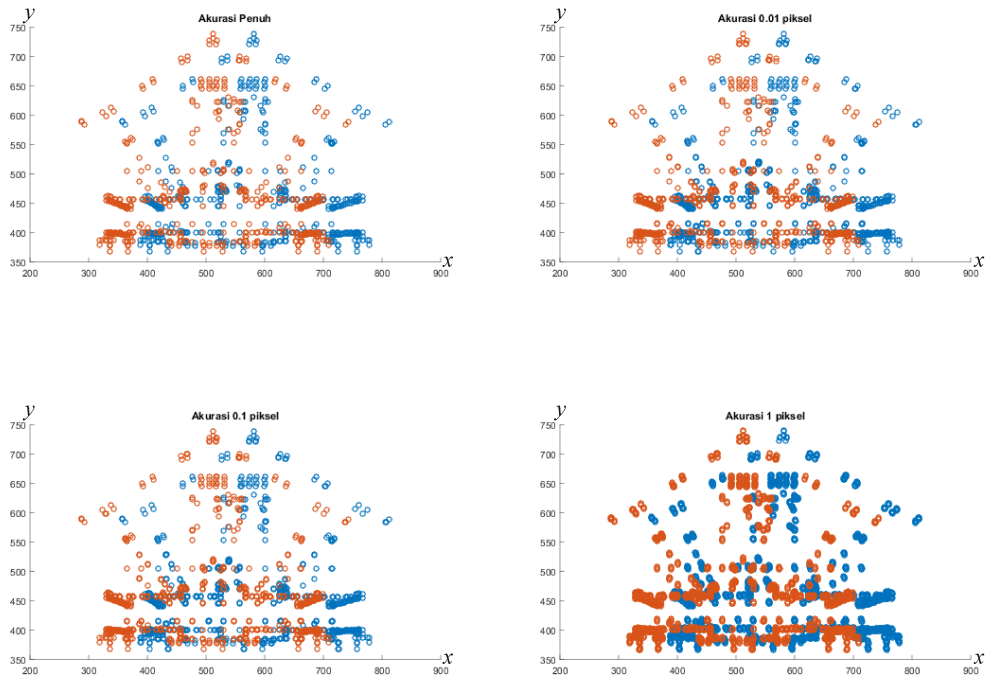
Gambar 4.46: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* keempat.

sama tersebar secara vertikal horizontal dengan persebaran yang cukup luas. Dihat dari bentuk persebarannya, sulit untuk memperkirakan keterpautan hasil estimasi *center of expansion*. Hal ini berpotensi mengakibatkan error pada hasil rekonstruksi yang cukup besar.

Dari hasil normalisasi ini, diduga titik *center of expansion* hasil estimasi masih memiliki perbedaan posisi yang cukup besar dan menyebabkan titik-titik yang berasal dari objek yang sama tidak terpetakan pada titik yang sama pada bidang normalisasi. Secara kasat mata, hasil normalisasi hanya memiliki perbedaan posisi tanpa perbedaan skala.

#### 4.4.2. Normalisasi menggunakan fundamental matriks

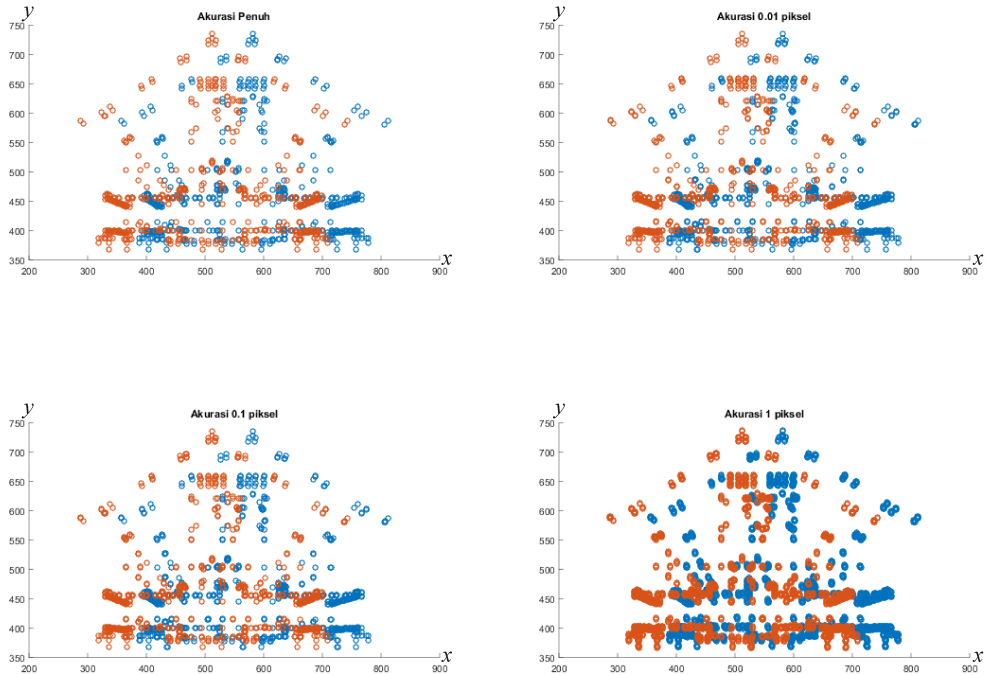
Pada metode ini faktor skala antar frame  $s_n$  yang digunakan adalah nilai faktor skala yang didapat menggunakan bantuan fundamental matriks. Untuk jenis data sintetis, data hasil proyeksi pada setiap kamera dinormalisasi secara terpisah dengan nilai *focal length* dan lokasi *center of expansion* hasil estimasi masing-masing. Normalisasi pada masing-masing simulasi akurasi proyeksi juga dilakukan secara terpisah. Penyajian data dikelompokkan berdasarkan je-



Gambar 4.47: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 1.

nis kamera dan skenario *zooming*. Bagian ini menyajikan hasil normalisasi dari titik-titik hasil proyeksi kamera pertama dan kedua pada masing-masing kamera stereo virtual di dalam satu bidang. Titik-titik berwarna biru merupakan hasil normalisasi titik-titik yang terproyeksi pada kamera pertama sedangkan titik-titik merah merupakan hasil normalisasi titik-titik yang terproyeksi pada kamera kedua. Hasil normalisasi disajikan pada Gambar 4.47 hingga Gambar 4.50.

Gambar 4.47 menunjukkan bahwa pada simulasi akurasi penuh hingga 0.1 piksel, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama berhasil ternormalisasi pada posisi yang relatif sama, namun seiring dengan menurunnya akurasi, persebaran hasil normalisasi dari titik-titik yang sama ikut membesar. Sedangkan untuk akurasi 1 piksel, dapat diamati bahwa titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada

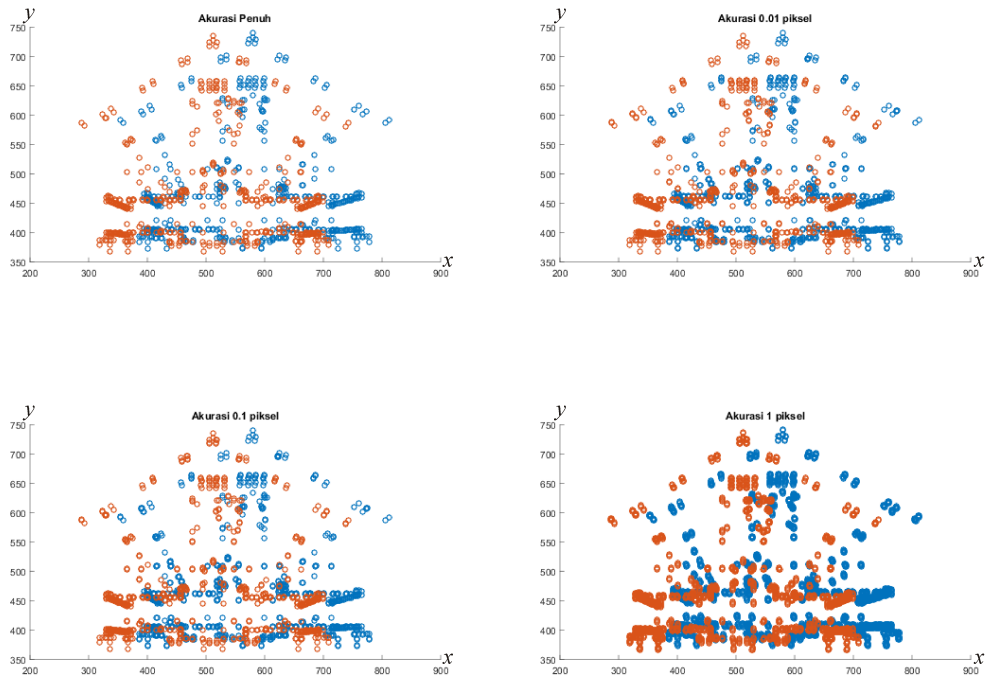


Gambar 4.48: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 2.

posisi-posisi yang cukup tersebar. Hal ini berarti bahwa untuk proyeksi pada kamera stereo virtual dengan *focal length* dan posisi *image center* yang sama, kesalahan korespondensi yang signifikan mulai terjadi pada tingkat akurasi 1 piksel. Hasil normalisasi pada Gambar 4.47 tidak menunjukkan perbedaan yang signifikan dengan Gambar 4.39.

Gambar 4.48 menunjukkan bahwa pada simulasi akurasi penuh hingga 0.1 piksel, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama berhasil ternormalisasi pada posisi yang relatif sama namun seiring dengan menurunnya akurasi, persebaran hasil normalisasi dari titik-titik yang sama ikut membesar. Sedangkan untuk akurasi 1 piksel, dapat diamati bahwa titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Hal ini berarti bahwa untuk proyeksi pada kamera stereo dengan posisi *image center* yang sama namun terdapat perbedaan  $f_x$  dan  $f_y$

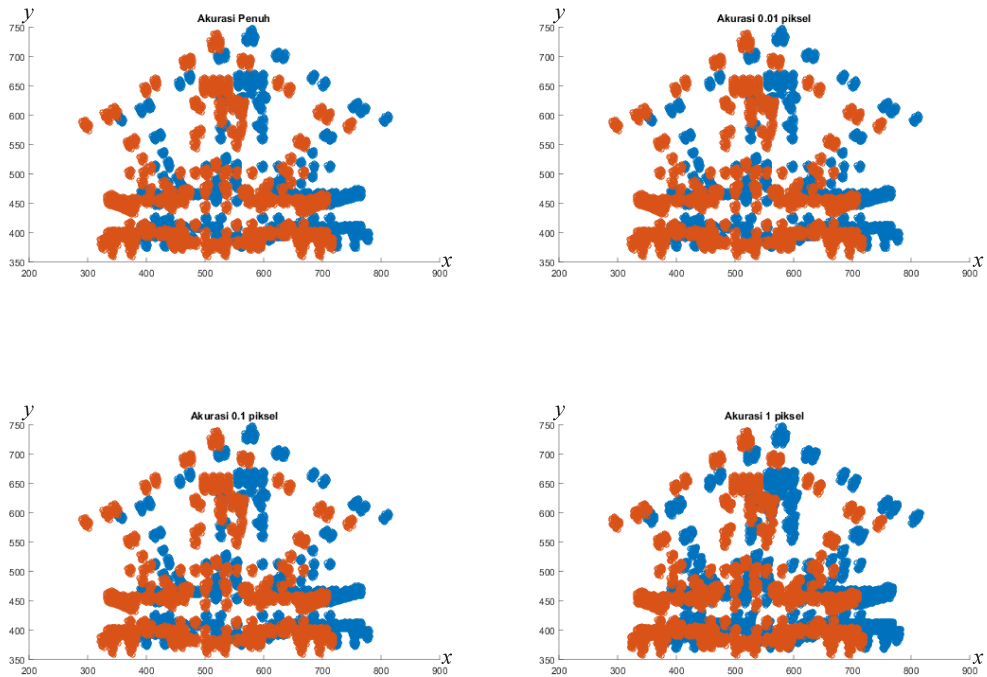




Gambar 4.49: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3.

pada masing-masing kamera, kesalahan korespondensi yang signifikan mulai terjadi pada tingkat akurasi 1 piksel. Hasil normalisasi pada Gambar 4.48 tidak menunjukkan perbedaan yang signifikan dengan Gambar 4.40.

Gambar 4.49 menunjukkan bahwa pada simulasi akurasi penuh hingga 0.1 piksel, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama berhasil ternormalisasi pada posisi yang relatif sama namun seiring dengan menurunnya akurasi, persebaran hasil normalisasi dari titik-titik yang sama ikut membesar. Sedangkan untuk akurasi 1 piksel, dapat diamati bahwa titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Hal ini berarti bahwa untuk proyeksi pada kamera stereo di mana masing-masing kamera memiliki parameter intrinsik yang berbeda, kesalahan korespondensi yang signifikan mulai terjadi pada tingkat akurasi 1 piksel. Hasil normalisasi pada Gambar 4.49 tidak menunjukkan perbedaan

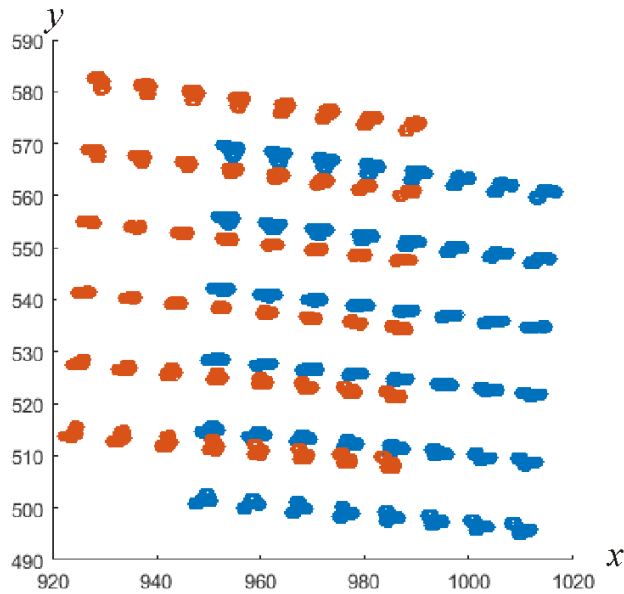


Gambar 4.50: Hasil normalisasi berbasis garis ekspansi dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

yang signifikan dengan Gambar 4.41.

Gambar 4.50 menunjukkan bahwa pada semua level simulasi akurasi, titik-titik dua dimensi pada citra yang merupakan hasil proyeksi dari titik tiga dimensi yang sama ternormalisasi pada posisi-posisi yang cukup tersebar. Jika dibandingkan dengan hasil-hasil pada normalisasi sebelumnya, perubahan posisi *image center* sangat menentukan akurasi dari korespondensi titik-titik dua dimensi hasil proyeksi. Hasil normalisasi pada Gambar 4.50 tidak menunjukkan perbedaan yang signifikan dengan Gambar 4.42.

Pada jenis data citra, bidang target normalisasi adalah frame pertama di mana level *zoom* pada frame pertama merupakan *zoom minimal*. Normalisasi menggunakan frame dengan *zoom* minimal karena pada level *zoom* tersebut parameter intrinsik dan ekstrinsik kamera diketahui dari hasil kalibrasi awal. Bagian ini menyajikan hasil normalisasi dari titik-titik hasil proyeksi kamera kiri dan kanan pada masing-masing kamera stereo di dalam satu bidang. Titik-

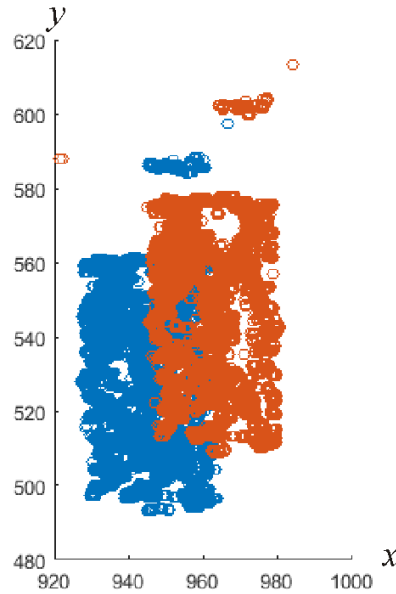


Gambar 4.51: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* pertama.

titik berwarna biru merupakan hasil normalisasi titik-titik objek atau lingkungan yang terproyeksi pada kamera kiri sedangkan titik-titik merah merupakan hasil normalisasi titik-titik yang terproyeksi pada kamera kanan. Penelitian ini melakukan normalisasi pada semua titik fitur yang ditemukan di dalam masing-masing *scene* pada seluruh frame. Pada *scene* pertama yang mengandung pola papan catur, penelitian ini hanya menggunakan titik-titik *cross corner* yang terdeteksi dan pada *scene* kedua hingga keempat seluruh titik fitur SURF digunakan. Hasil normalisasi ditampilkan pada Gambar 4.52 hingga Gambar 4.54.

Gambar 4.51 menunjukkan bahwa data hasil normalisasi dari objek yang sama tersebar secara vertikal horizontal dengan persebaran yang signifikan. Dilhat dari persebarannya yang lebih kecil secara vertikal maupun horizontal, dapat disimpulkan bahwa estimasi *center of expansion* dengan memanfaatkan fundamental matriks lebih mendekati posisi *center of expansion* asli jika dibandingkan dengan estimasi berbasis garis ekspansi.

Gambar 4.52 kumpulan titik hasil normalisasi memberikan bentuk yang menyerupai persegi panjang. Hasil normalisasi dari masing-masing proyeksi



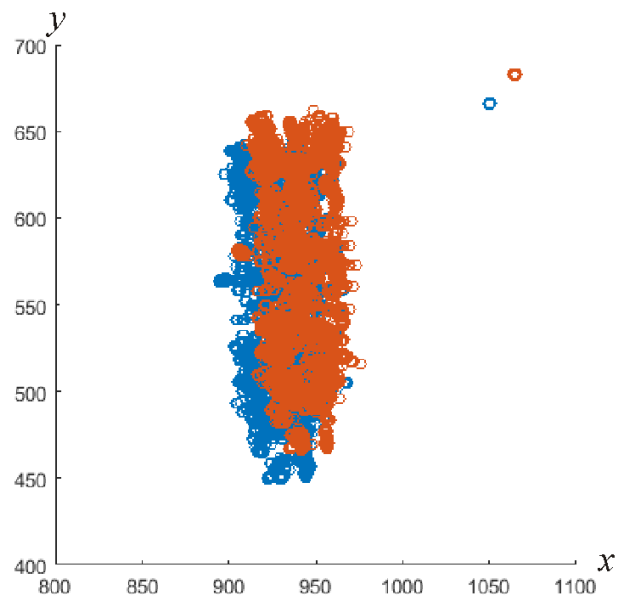
Gambar 4.52: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* kedua.

kamera tampak mirip dengan hasil pada Gambar 4.44 namun dengan *disparity* antar citra yang lebih besar.

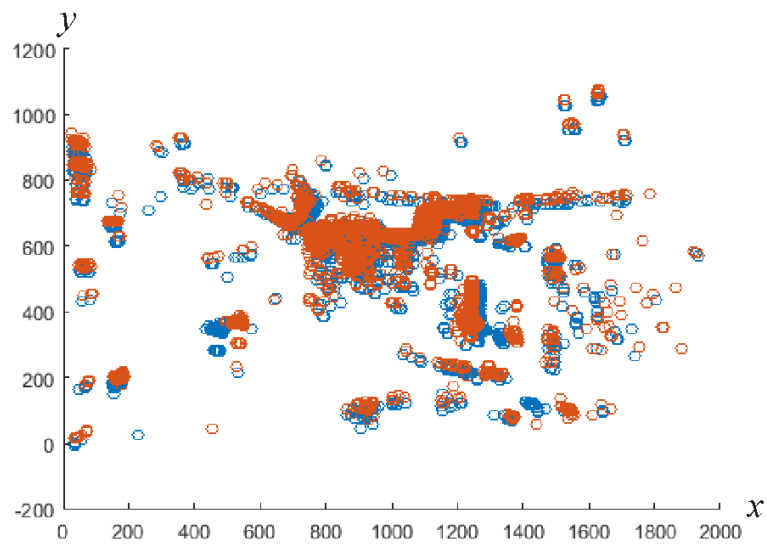
Gambar 4.53 kumpulan titik hasil normalisasi memberikan bentuk yang menyerupai botol terbalik. Hasil normalisasi ini tampak lebih ramping jika dibandingkan dengan Gambar 4.45. Hal ini menjadi indikasi bahwa persebaran titik hasil normalisasi untuk objek yang sama memiliki persebaran secara horizontal yang lebih kecil.

Pada Gambar 4.54 tampak bahwa data hasil normalisasi dari objek yang sama tersebar secara vertikal horizontal dengan persebaran yang cukup luas. Dilhat dari bentuk penyebarannya, sulit untuk memperkirakan keterpautan hasil estimasi *center of expansion*. Hal ini berpotensi mengakibatkan error pada hasil rekonstruksi yang cukup besar. Tidak tampak perbedaan distribusi titik yang signifikan jika dibandingkan dengan Gambar 4.54.

Dari hasil normalisasi ini, diduga titik *center of expansion* hasil estimasi masih memiliki perbedaan posisi yang dan menyebabkan titik-titik yang berasal dari objek yang sama tidak terpetakan pada titik yang sama pada bidang normalisasi. Secara umum, distribusi titik-titik normalisasi dari objek yang



Gambar 4.53: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* ketiga.



Gambar 4.54: Hasil normalisasi berbasis garis ekspansi dari data citra *scene* keempat.

sama lebih kecil jika dibandingkan dengan metode berbasis garis-garis ekspansi.

#### 4.5. Rekonstruksi Tiga Dimensi

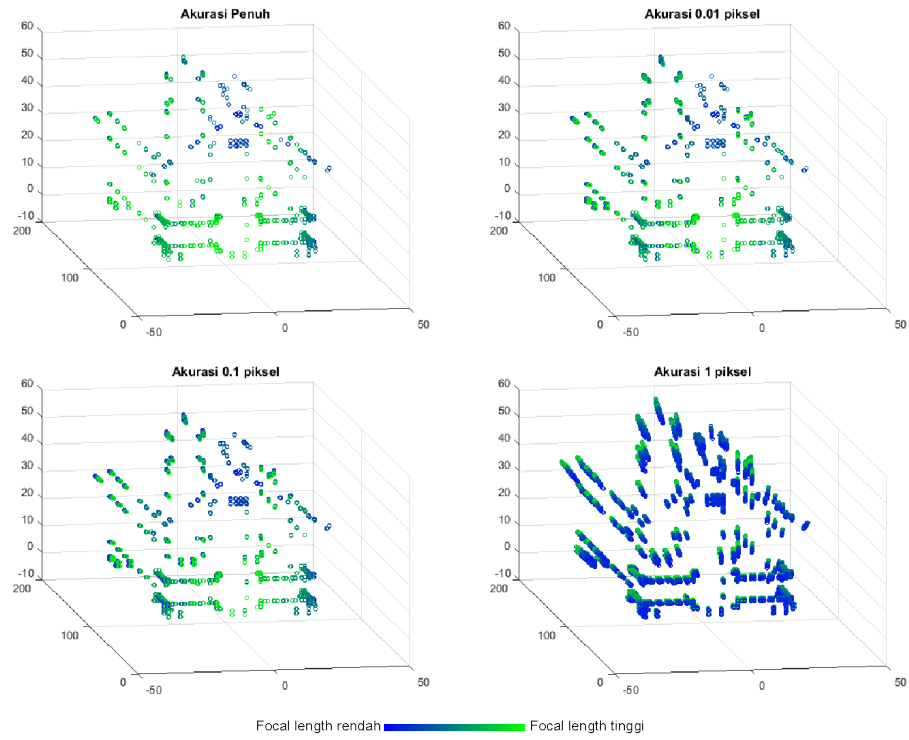
Proses rekonstruksi ini membangun titik-titik tiga dimensi dari titik-titik proyeksi yang telah dinormalisasi baik dari data sintetis maupun data citra. Parameter intrinsik dan ekstrinsik dari kamera yang digunakan adalah parameter kamera pada frame atau *sequence* pertama mengikuti bidang yang menjadi target normalisasi. Pembahasan pada subbab ini dibagi menjadi dua berdasarkan metode yang digunakan untuk estimasi *focal length* dan *center of expansion* yaitu metode yang tidak menggunakan fundamental matriks dan metode yang memanfaatkan fundamental matriks. Pada masing-masing metode, pembahasan diawali dengan jenis data sintetis dan dilanjutkan dengan jenis data citra.

##### 4.5.1. Rekonstruksi tanpa fundamental matriks

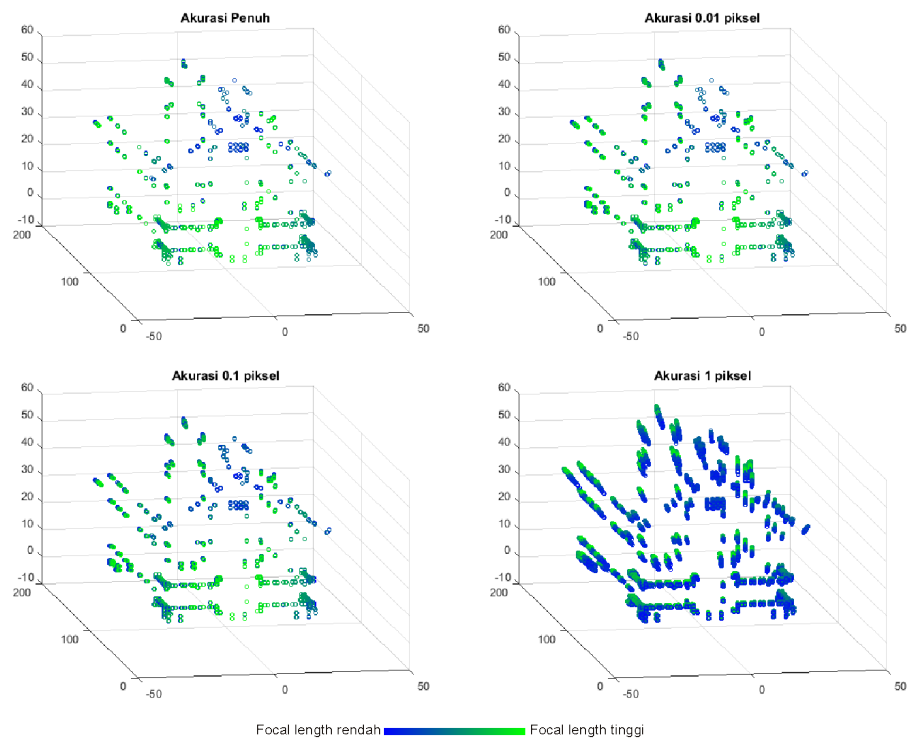
Rekonstruksi jenis ini adalah rekonstruksi yang menggunakan jarak titik sebagai basis estimasi *focal length* dan garis ekspansi sebagai basis estimasi *center of expansion*. Hasil rekonstruksi dari data sintetis dan data citra disajikan pada Gambar 4.55 hingga Gambar 4.62.

Pada Gambar 4.55 tampak bahwa pada simulasi akurasi penuh hingga 0.1 piksel tidak terlihat perbedaan hasil transformasi yang signifikan pada seluruh frame. Pada simulasi akurasi 1 piksel mulai tampak distorsi pada hasil rekonstruksi. Hal ini menunjukkan bahwa dalam kondisi parameter intrinsik kamera yang identik serta  $f_x$  dan  $f_y$  yang sama, estimasi *focal length* dan *center of expansion* untuk normalisasi memberikan hasil rekonstruksi yang akurat selama ketelitian korespondensi tidak lebih rendah dari 0.1 piksel.

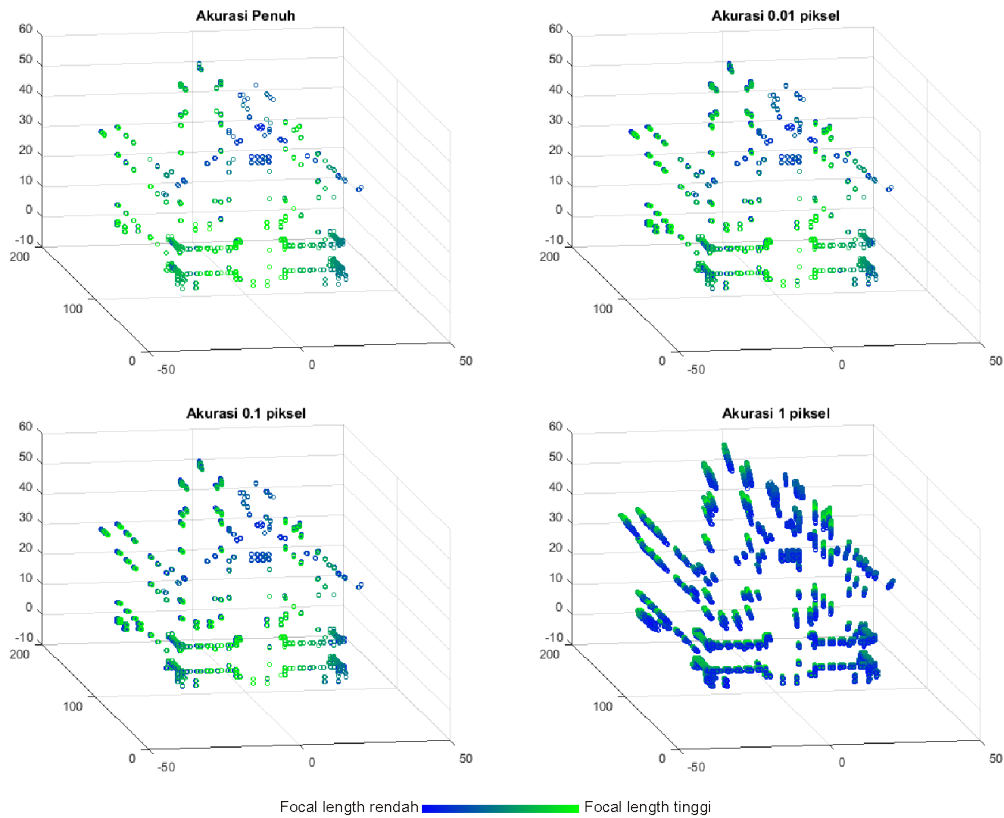
Pada Gambar 4.56 tampak bahwa pada simulasi akurasi penuh hingga 0.1 piksel tidak terlihat perbedaan hasil transformasi yang signifikan pada seluruh frame. Pada simulasi akurasi 1 piksel mulai tampak distorsi pada hasil rekonstruksi. Hal ini menunjukkan bahwa dalam kondisi parameter intrinsik



Gambar 4.55: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1.



Gambar 4.56: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2.



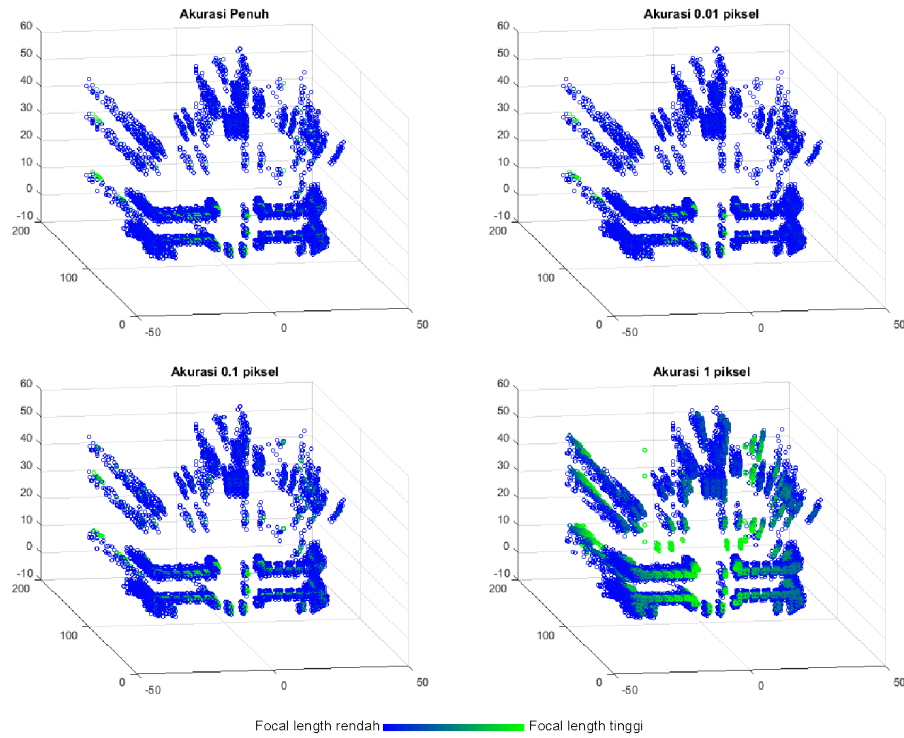
Gambar 4.57: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintesis menggunakan kamera stereo virtual 3.

kamera yang identik serta  $f_x$  dan  $f_y$  yang berbeda, estimasi *focal length* dan *center of expansion* untuk normalisasi memberikan hasil rekonstruksi yang akurat selama ketelitian korespondensi tidak lebih rendah dari 0.1 piksel.

Pada Gambar 4.57 tampak bahwa pada simulasi akurasi penuh hingga 0.1 piksel tidak terlihat perbedaan hasil transformasi yang signifikan pada seluruh frame. Pada simulasi akurasi 1 piksel mulai tampak distorsi pada hasil rekonstruksi. Hal ini menunjukkan bahwa dalam kondisi parameter intrinsik kamera yang berbeda, estimasi *focal length* dan *center of expansion* untuk normalisasi memberikan hasil rekonstruksi yang akurat selama ketelitian korespondensi tidak lebih rendah dari 0.1 piksel.

Pada Gambar 4.58 tampak bahwa pada seluruh simulasi akurasi korespondensi memberikan hasil rekonstruksi yang mengalami distorsi. Hal ini menunjukkan bahwa posisi *image center* relatif terhadap titik *center of expansion* terus berubah selama proses *zooming*. Perubahan *center of expansion* yang



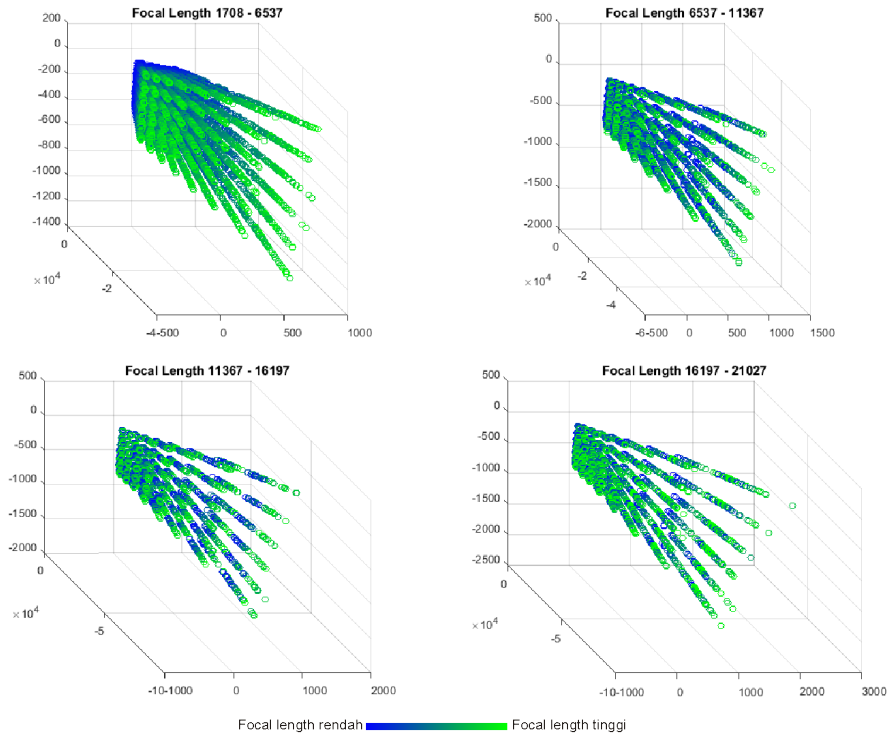


Gambar 4.58: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

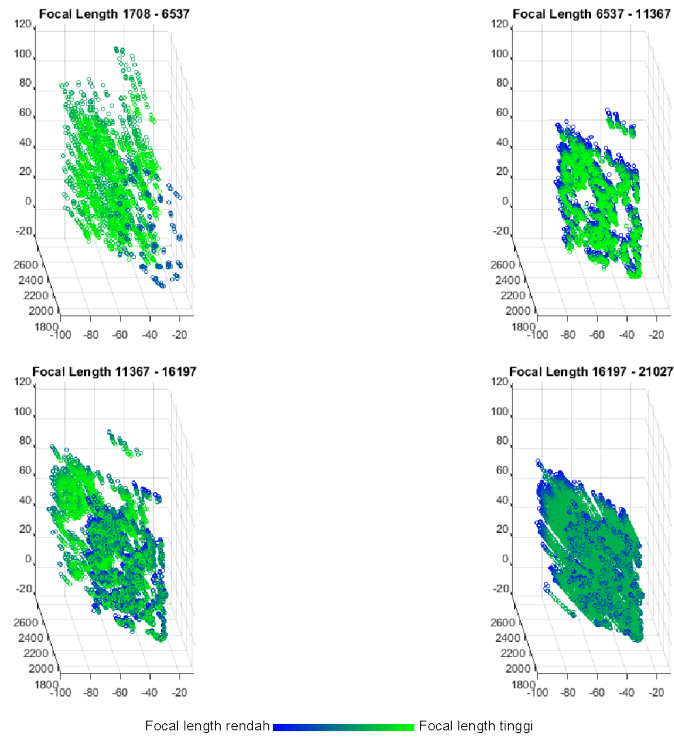
tidak sepenuhnya tergantung pada proses *scaling* mengakibatkan perubahan titik tidak ternormalisasi secara tepat sehingga hasil rekonstruksi untuk titik yang sama lebih tersebar.

Pada Gambar 4.59 tampak bahwa hasil rekonstruksi berhasil membentuk pola papan catur. Hasil rekonstruksi mengalami transformasi skala pada sumbu  $x$ ,  $y$  dan  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normalisasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses zooming.

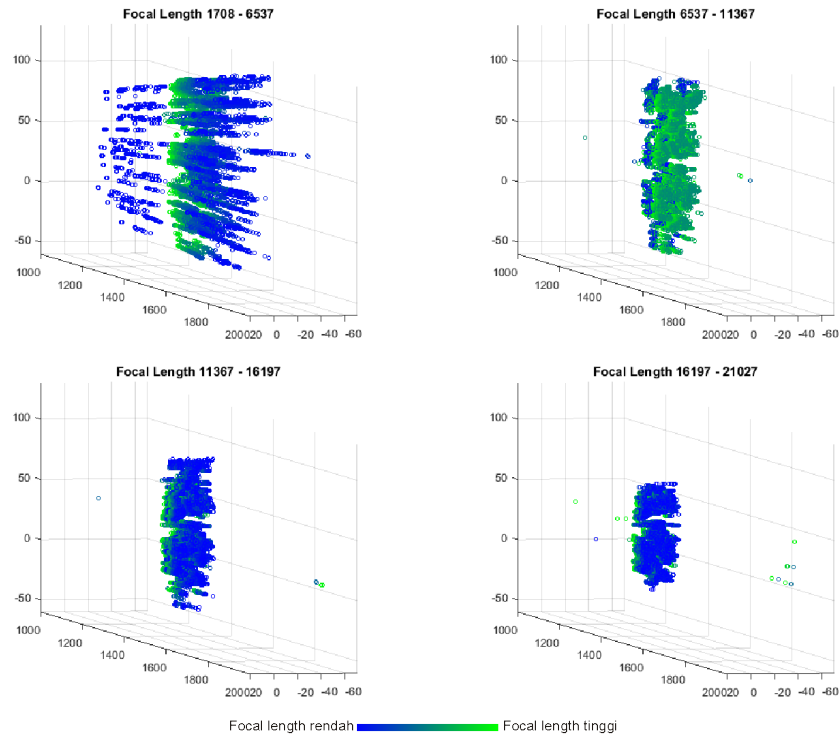
Pada Gambar 4.60 tampak bahwa hasil rekonstruksi tampak membentuk siluet persegi panjang yang mengalami transformasi *skew* vertikal. Hasil rekonstruksi mengalami transformasi skala pada sumbu  $x$ ,  $y$  dan  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normal-



Gambar 4.59: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra *scene* pertama.



Gambar 4.60: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra *scene* kedua.

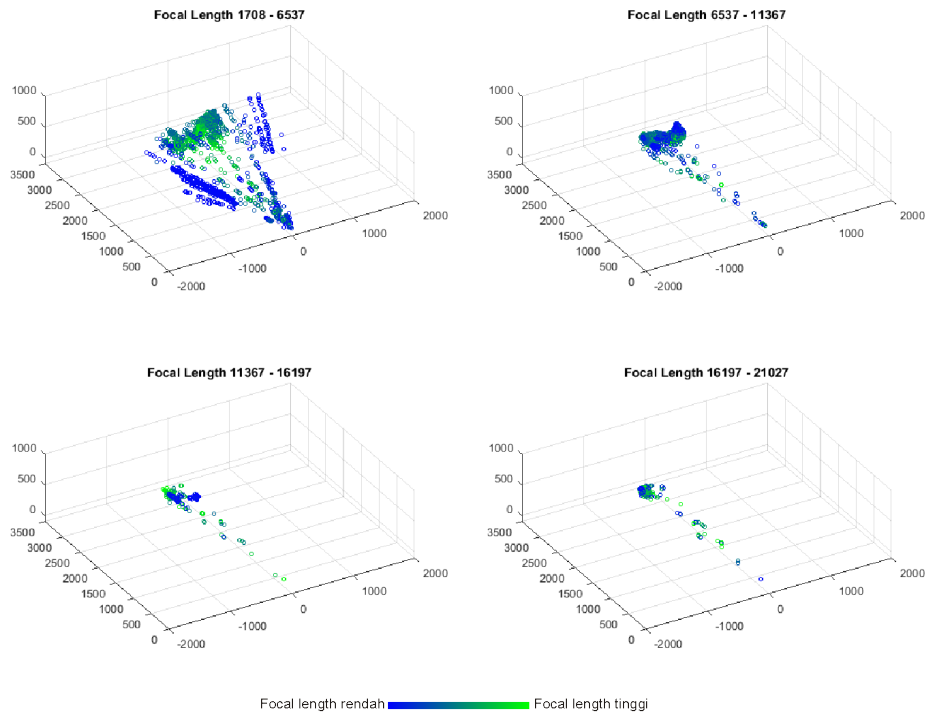


Gambar 4.61: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra *scene* ketiga.

isasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses zooming.

Pada Gambar 4.61 tampak bahwa hasil rekonstruksi tampak membentuk siluet botol terbalik. Hasil rekonstruksi mengalami transformasi skala pada sumbu  $x$ ,  $y$  dan  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normalisasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses zooming.

Pada Gambar 4.62 tampak bahwa hasil rekonstruksi tampak membentuk beberapa objek di dalam lingkungan outdoor. Hasil rekonstruksi mengalami transformasi skala pada sumbu  $x$ ,  $y$  dan  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normalisasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses zooming.



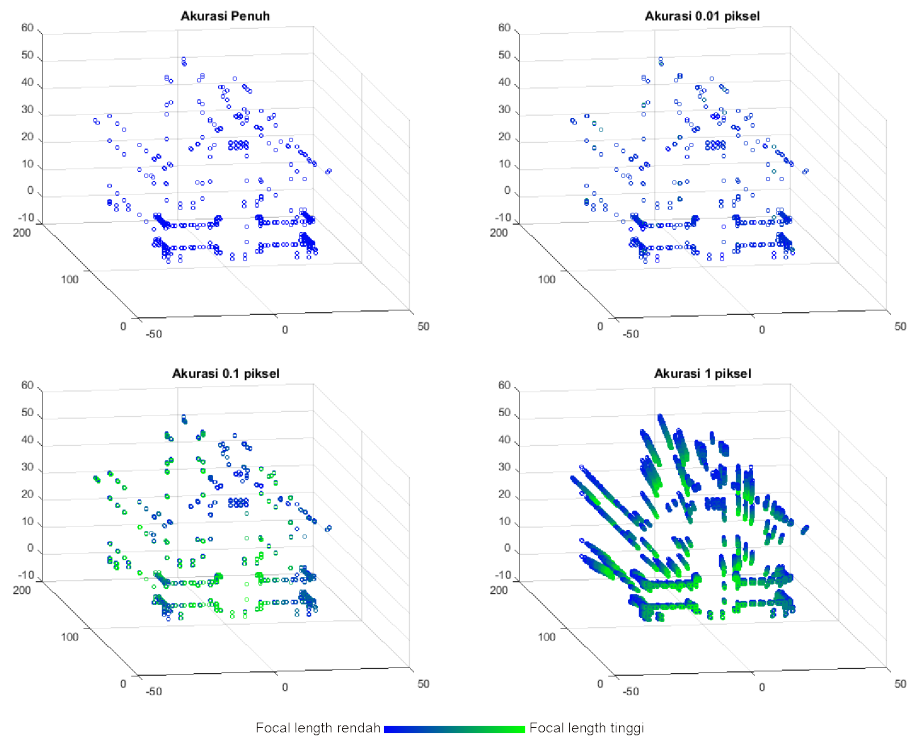
Gambar 4.62: Hasil rekonstruksi menggunakan estimasi tanpa fundamental matriks dari data citra *scene* keempat.

#### 4.5.2. Rekonstruksi memanfaatkan fundamental matriks

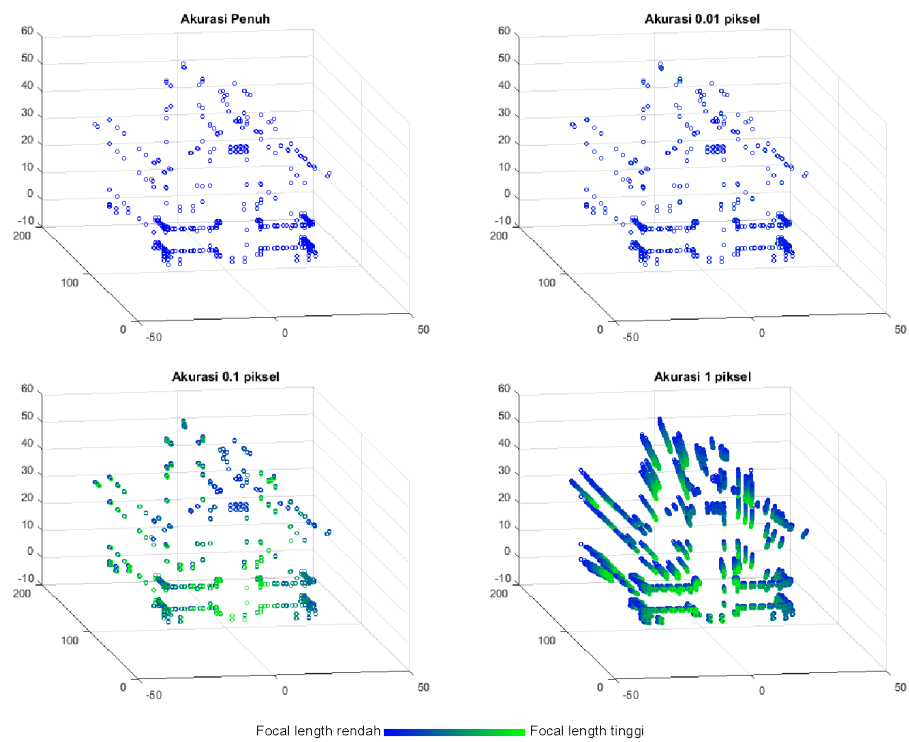
Rekonstruksi jenis ini adalah rekonstruksi yang memanfaatkan fundamental sebagai basis estimasi *focal length* dan estimasi *center of expansion*. Hasil rekonstruksi dari data sintesis dan data citra disajikan pada Gambar 4.63 hingga Gambar 4.70.

Pada Gambar 4.63 tampak bahwa pada simulasi akurasi penuh hingga 0.1 piksel tidak terlihat perbedaan hasil transformasi yang signifikan pada seluruh frame. Pada simulasi akurasi 1 piksel mulai tampak distorsi pada hasil rekonstruksi. Hal ini menunjukkan bahwa dalam kondisi parameter intrinsik kamera yang identik serta  $f_x$  dan  $f_y$  yang sama, estimasi *focal length* dan *center of expansion* untuk normalisasi memberikan hasil rekonstruksi yang akurat selama ketelitian korespondensi tidak lebih rendah dari 0.1 piksel. Perbedaan hasil rekonstruksi tidak tampak jika dibandingkan dengan Gambar 4.55.

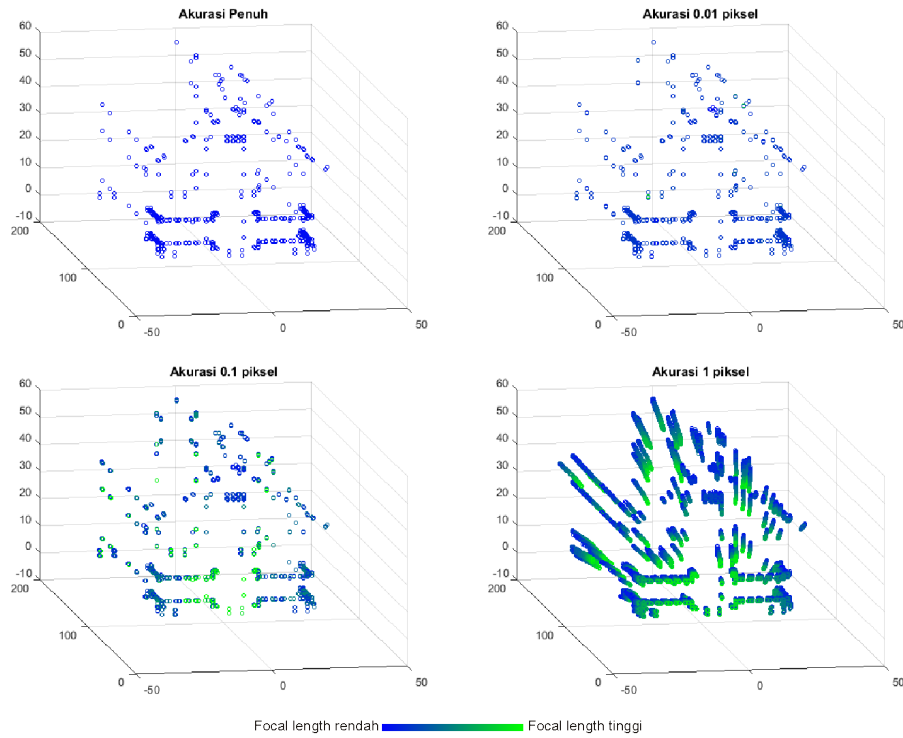
Pada Gambar 4.64 tampak bahwa pada simulasi akurasi penuh hingga 0.1 piksel tidak terlihat perbedaan hasil transformasi yang signifikan pada seluruh



Gambar 4.63: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 1.



Gambar 4.64: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 2.

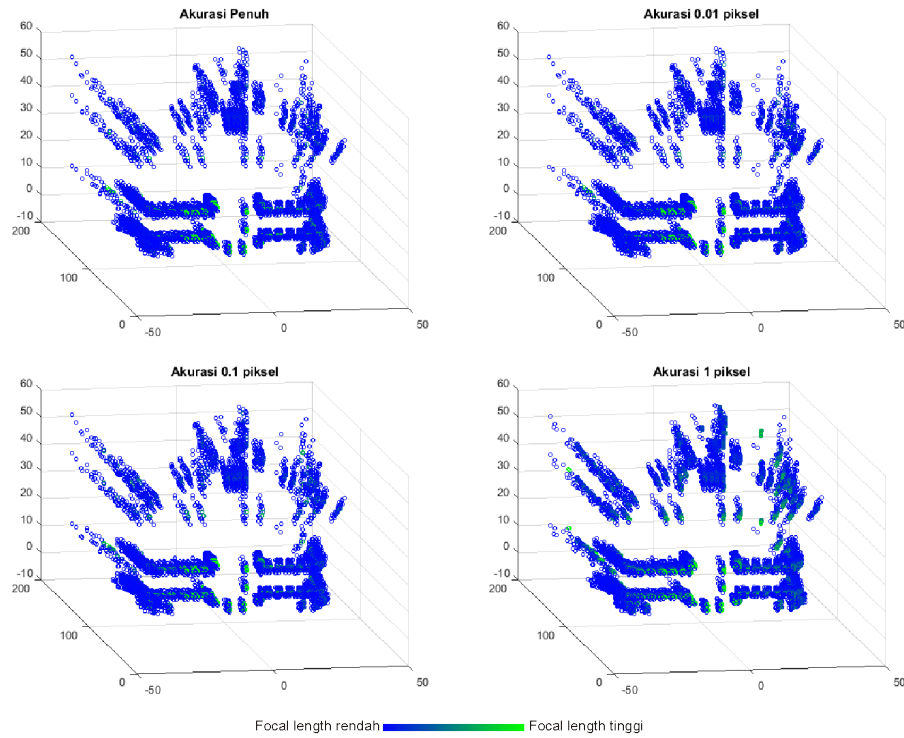


Gambar 4.65: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3.

frame. Pada simulasi akurasi 1 piksel mulai tampak distorsi pada hasil rekonstruksi. Hal ini menunjukkan bahwa dalam kondisi parameter intrinsik kamera yang identik serta  $f_x$  dan  $f_y$  yang berbeda, estimasi *focal length* dan *center of expansion* untuk normalisasi memberikan hasil rekonstruksi yang akurat selama ketelitian korespondensi tidak lebih rendah dari 0.1 piksel. Perbedaan hasil rekonstruksi tidak tampak jika dibandingkan dengan Gambar 4.56.

Pada Gambar 4.65 tampak bahwa pada simulasi akurasi penuh hingga 0.1 piksel tidak terlihat perbedaan hasil transformasi yang signifikan pada seluruh frame. Pada simulasi akurasi 1 piksel mulai tampak distorsi pada hasil rekonstruksi. Hal ini menunjukkan bahwa dalam kondisi parameter intrinsik kamera yang berbeda, estimasi *focal length* dan *center of expansion* untuk normalisasi memberikan hasil rekonstruksi yang akurat selama ketelitian korespondensi tidak lebih rendah dari 0.1 piksel. Perbedaan hasil rekonstruksi tidak tampak jika dibandingkan dengan Gambar 4.57.

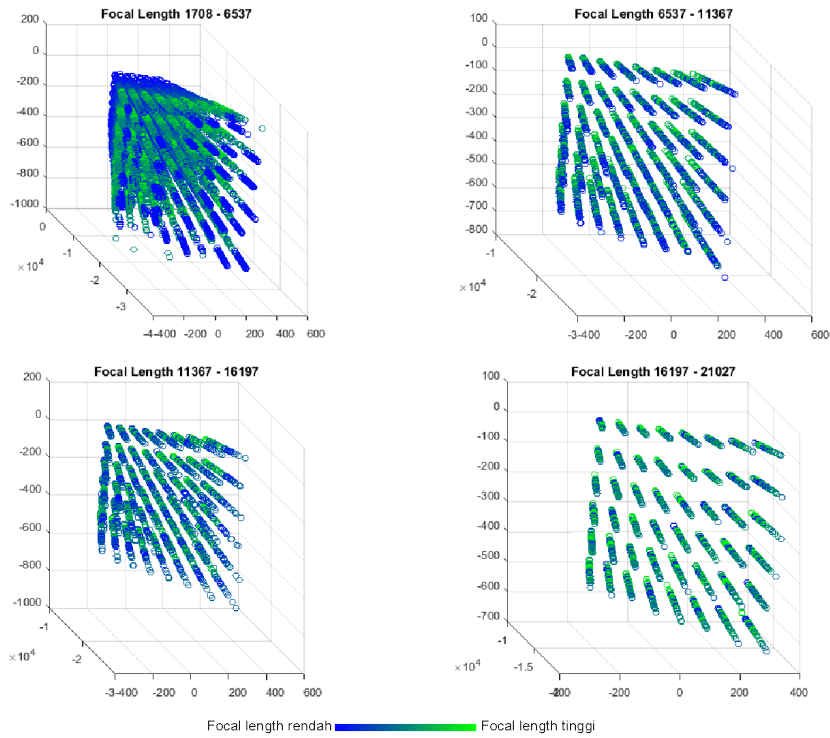
Pada Gambar 4.66 tampak bahwa pada seluruh simulasi akurasi korespon-



Gambar 4.66: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data sintetis menggunakan kamera stereo virtual 3 dengan *image center* dinamis.

densi memberikan hasil rekonstruksi yang mengalami distorsi. Hal ini menunjukkan bahwa posisi *image center* relatif terhadap titik *center of expansion* terus berubah selama proses *zooming*. Perubahan *center of expansion* yang tidak sepenuhnya tergantung pada proses *scaling* mengakibatkan perubahan titik tidak ternormalisasi secara tepat sehingga hasil rekonstruksi untuk titik yang sama lebih tersebar. Perbedaan hasil rekonstruksi tidak tampak jika dibandingkan dengan Gambar 4.58.

Pada Gambar 4.67 tampak bahwa hasil rekonstruksi berhasil membentuk pola papan catur. Hasil rekonstruksi mengalami transformasi skala pada sumbu  $x$ ,  $y$  dan  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normalisasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses *zooming*. Jika dibandingkan dengan 4.59, hasil rekonstruksi pada titik yang sama memiliki persebaran yang



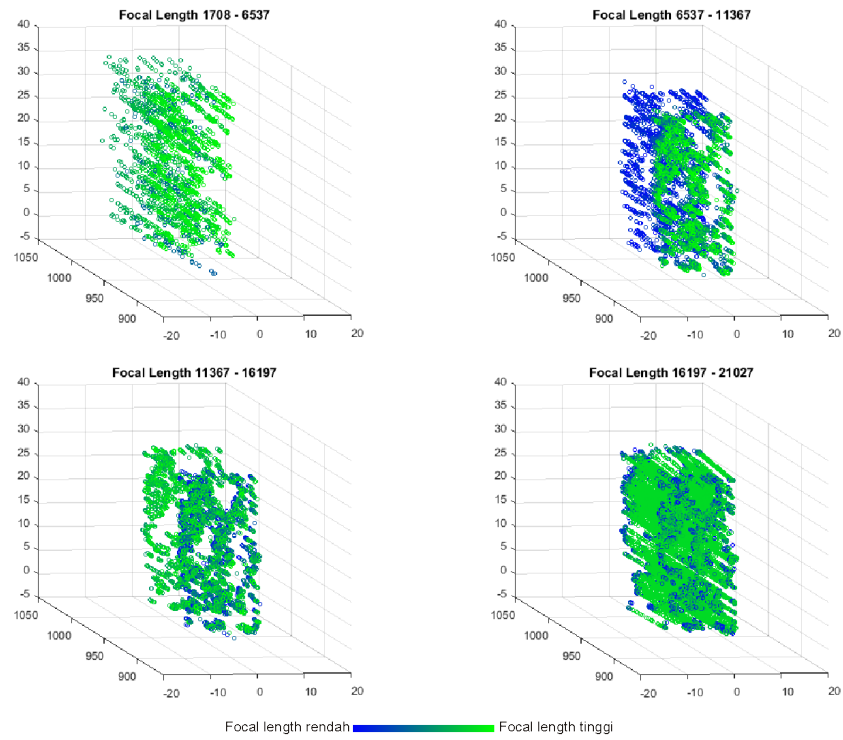
Gambar 4.67: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra *scene* pertama.

lebih kecil. Hal ini mengindikasikan bahwa metode estimasi untuk normalisasi menggunakan fundamental matriks mampu menjaga korespondensi titik lebih baik.

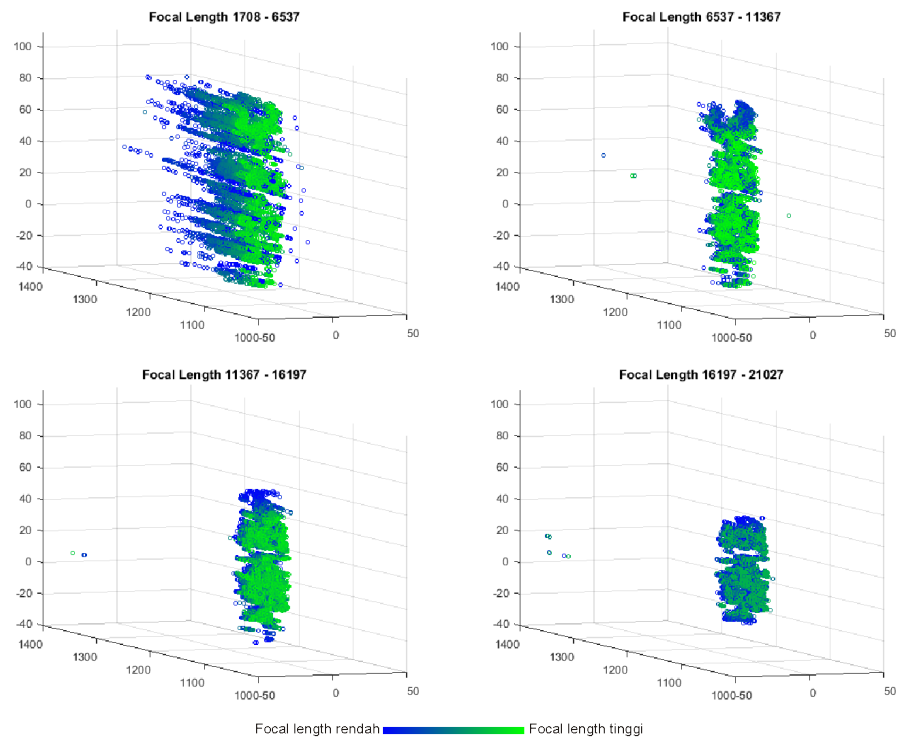
Pada Gambar 4.68 tampak bahwa hasil rekonstruksi tampak membentuk persegi panjang sesuai dengan objek. Hasil rekonstruksi mengalami transformasi skala atau translasi yang mayoritas sejajar dengan sumbu  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normalisasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses zooming. Jika dibandingkan dengan 4.61, hasil rekonstruksi objek lebih mendekati bentuk fisik objek asli. Hal ini mengindikasikan bahwa metode estimasi untuk normalisasi menggunakan fundamental matriks mampu memberikan bentuk rekonstruksi yang lebih akurat.

Pada Gambar 4.69 tampak bahwa hasil rekonstruksi tampak membentuk siluet botol terbalik. Hasil rekonstruksi mengalami transformasi skala atau





Gambar 4.68: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra *scene* kedua.



Gambar 4.69: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra *scene* ketiga.

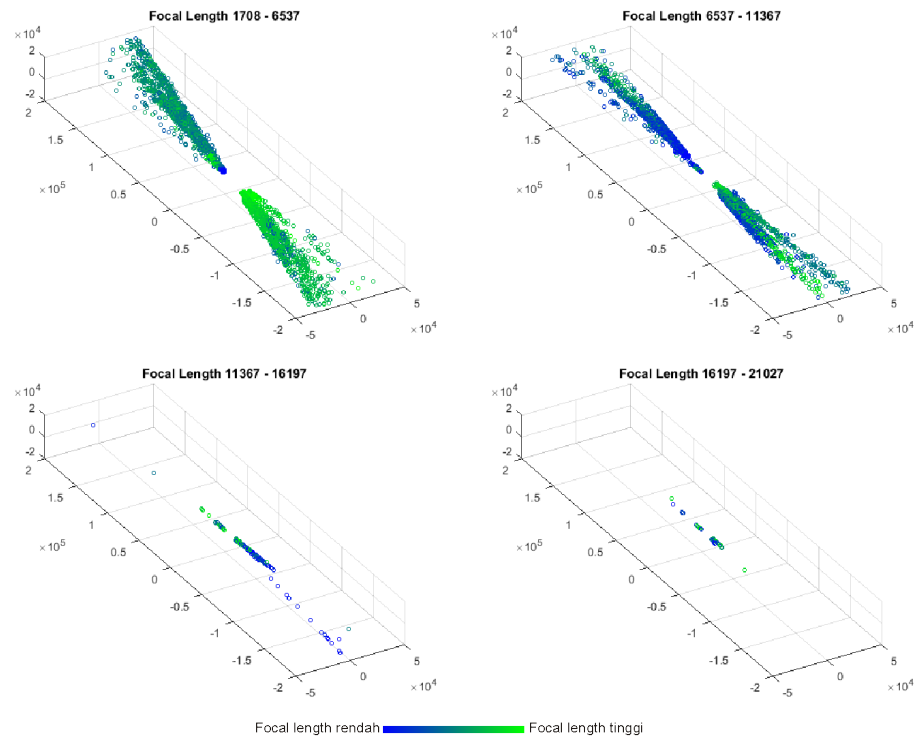
translasi yang mayoritas sejajar dengan sumbu  $z$  di setiap frame dengan *focal length* yang berbeda. Hal ini dapat disebabkan oleh tingkat akurasi korespondensi antar citra yang terlalu dan/atau kesalahan normalisasi yang disebabkan oleh perubahan posisi *image center* pada setiap frame yang mengalami proses zooming. Jika dibandingkan dengan 4.61, hasil rekonstruksi pada titik yang sama memiliki persebaran yang lebih kecil. Hal ini mengindikasikan bahwa metode estimasi untuk normalisasi menggunakan fundamental matriks mampu menjaga korespondensi titik lebih baik.

Pada Gambar 4.70 tampak distribusi titik-titik hasil rekonstruksi membentuk pola seperti jam pasir dengan titik  $(0, 0, 0)$  sebagai pusat. Hal ini disebabkan oleh kesalahan korespondensi titik setelah mengalami proses normalisasi di mana terdapat perbedaan koordinat pada sumbu  $x$  dari sebagian titik-titik yang berkorespondensi antara citra kiri dan kanan bernilai positif dan sebagian yang lain bernilai negatif. Nilai  $z$  dari hasil rekonstruksi bernilai positif jika nilai  $x$  dari titik di kamera kiri lebih besar daripada nilai  $x$  dari titik yang berkorespondensi di kamera kanan. Jika terjadi sebaliknya, hasil proyeksi balik dari kedua titik akan berpotongan di  $z$  negatif. Kesalahan normalisasi ini dapat disebabkan kesalahan korespondensi titik pada citra secara horizontal dikarenakan jenis data citra yang mengandung pengulangan pola tertentu secara horizontal sehingga menyebabkan kesalahan estimasi posisi *center of expansion*.

Hasil yang tampak pada tahap rekonstruksi ini adalah rekonstruksi objek dengan estimasi menggunakan fundamental matriks memberikan hasil yang setara pada jenis data sintetis. Pada jenis data citra, hasil rekonstruksi cenderung lebih sebangun dengan objek kecuali pada data *scene* empat.

#### 4.6. Evaluasi Hasil

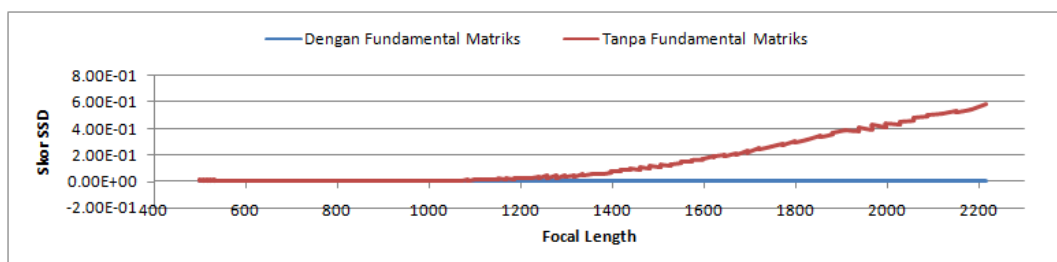
Bagian ini membagi pembahasan berdasarkan jenis data yang digunakan yaitu data sintetis dan data citra. Evaluasi dilakukan dengan membandingkan dua metode yang berbeda pada masing-masing jenis data.



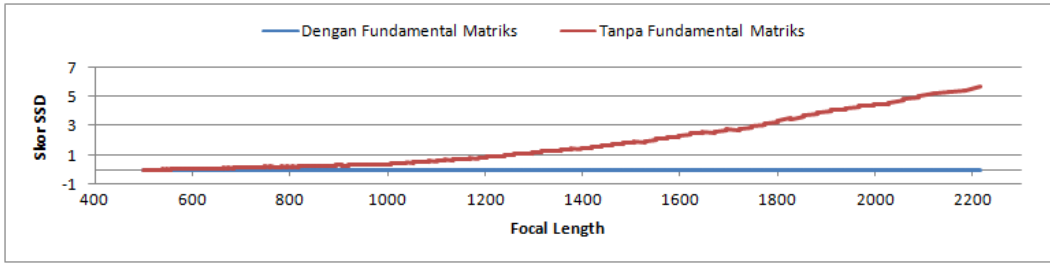
Gambar 4.70: Hasil rekonstruksi dengan estimasi memanfaatkan fundamental matriks dari data citra *scene* keempat.

#### 4.6.1. Evaluasi pada data sintetis

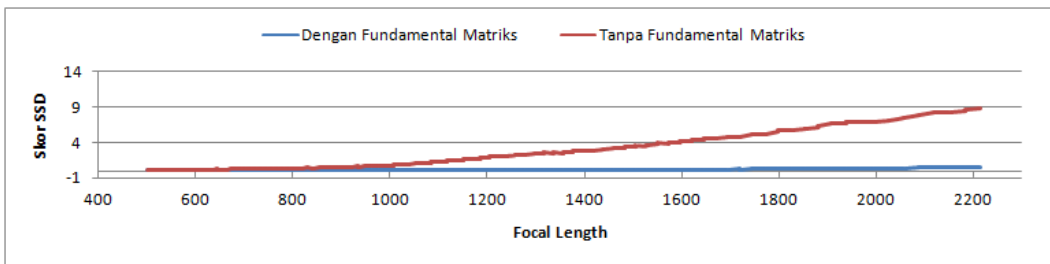
Pada jenis data sintetis penelitian ini melakukan perbandingan hasil evaluasi rekonstruksi dari kedua jenis metode pada masing-masing kamera. Perbandingan evaluasi MSE dan Pearson antara kedua metode dapat dilihat pada Gambar 4.71 hingga Gambar 4.102.



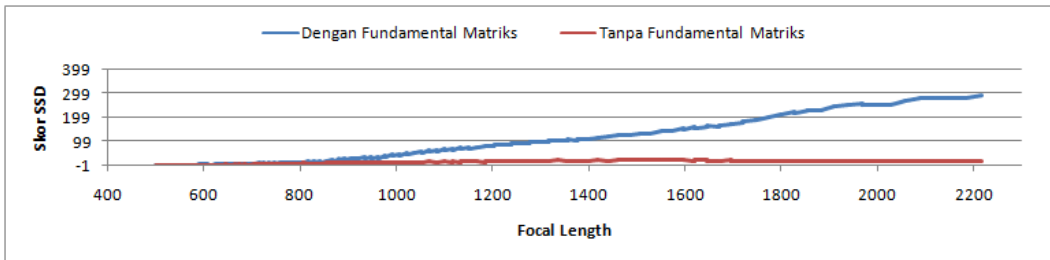
Gambar 4.71: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi penuh.



Gambar 4.72: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.01 piksel.

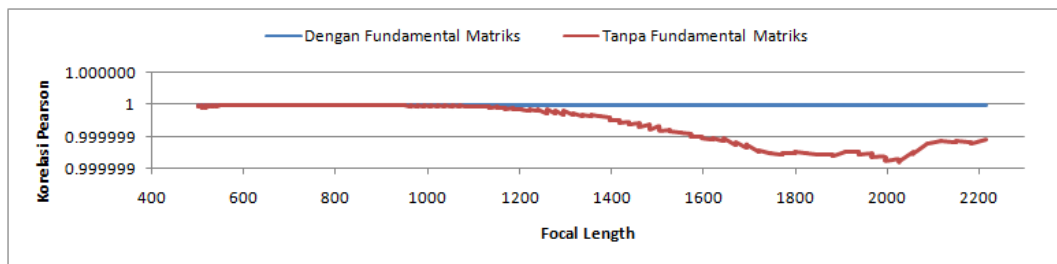


Gambar 4.73: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.1 piksel.

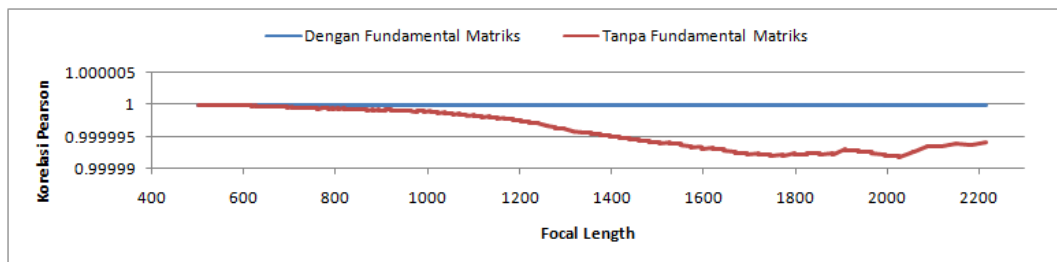


Gambar 4.74: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 1 piksel.

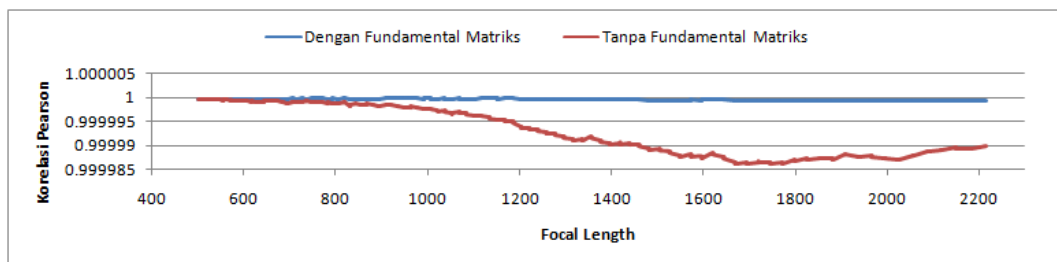
Gambar 4.71 hingga Gambar 4.74 menunjukkan perbandingan MSE hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual 1. Pada 4.71 hingga Gambar 4.73 menunjukkan bahwa hasil rekonstruksi yang melibatkan estimasi menggunakan fundamental matriks memberikan hasil yang lebih mendekati *ground truth* hingga pada akurasi 0.1 piksel. Pada Gambar 4.74, hasil rekonstruksi yang tidak melibatkan fundamental matriks memberikan hasil yang lebih baik meskipun Gambar 4.55 dan 4.63 menunjukkan hasil rekonstruksi mengalami distorsi.



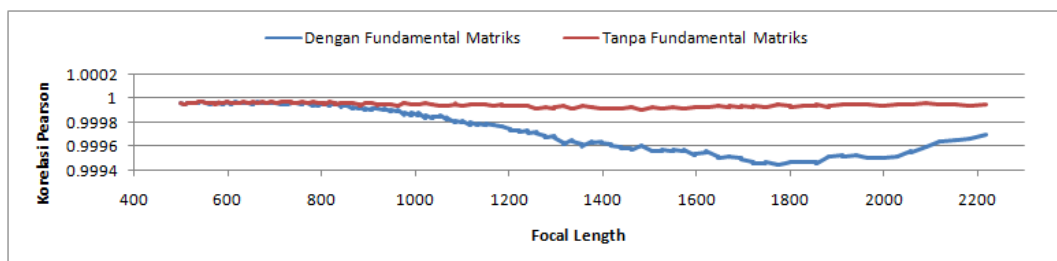
Gambar 4.75: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi penuh.



Gambar 4.76: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.01 piksel.



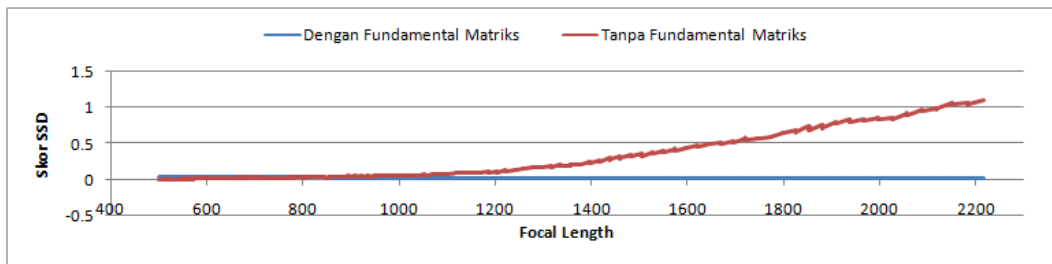
Gambar 4.77: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 0.1 piksel.



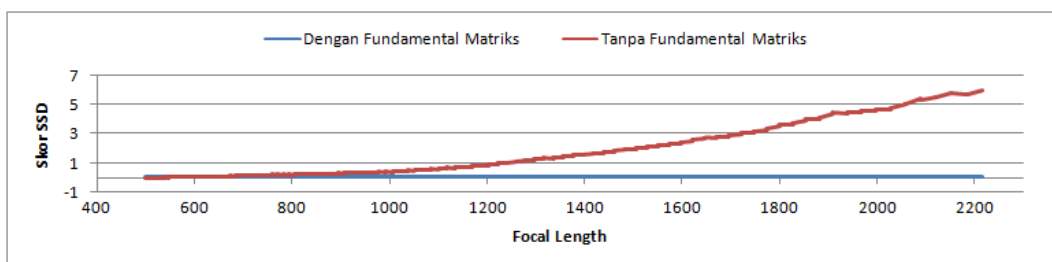
Gambar 4.78: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 1 pada simulasi akurasi 1 piksel.

Gambar 4.75 hingga Gambar 4.78 menunjukkan perbandingan Pearson hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo

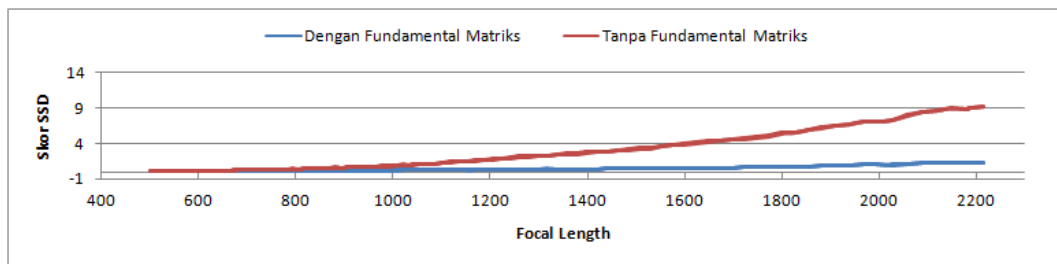
virtual 1. Skor evaluasi Pearson memberikan indikasi apakah hasil rekonstruksi masih memiliki bentuk yang serupa dengan *ground truth* meskipun memiliki ukuran yang berbeda. Jika skor Pearson semakin mendekati 1 maka hasil rekonstruksi semakin mendekati bentuk *ground truth*. Pada 4.75 hingga Gambar 4.77 menunjukkan bahwa hasil rekonstruksi yang melibatkan estimasi menggunakan fundamental matriks memberikan hasil yang lebih sebangun dengan *ground truth* hingga pada akurasi 0.1 piksel. Pada Gambar 4.78, hasil rekonstruksi yang tidak melibatkan fundamental matriks memberikan hasil yang lebih sebangun meskipun Gambar 4.55 dan 4.63 menunjukkan hasil rekonstruksi mengalami distorsi.



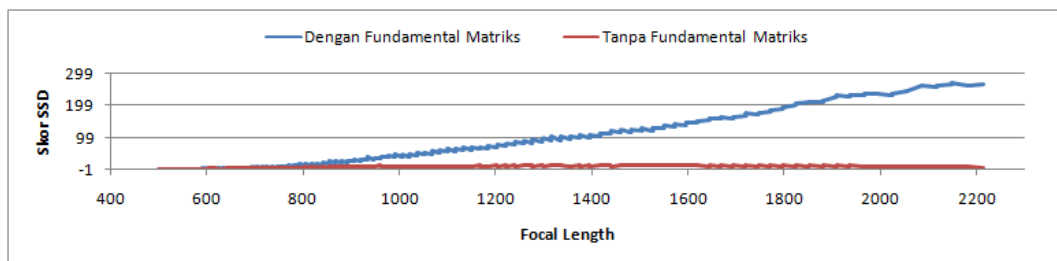
Gambar 4.79: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi penuh.



Gambar 4.80: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.01 piksel.

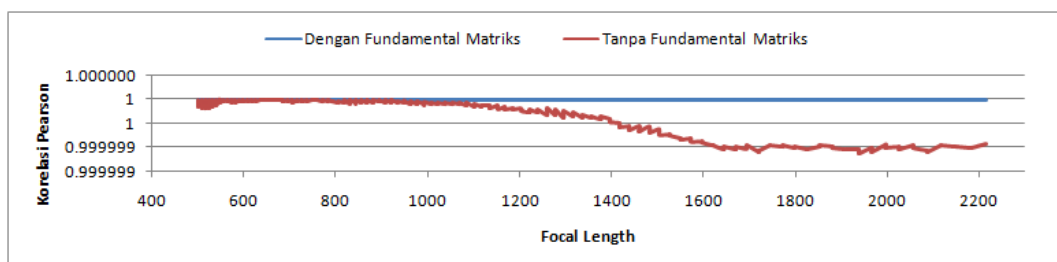


Gambar 4.81: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.1 piksel.

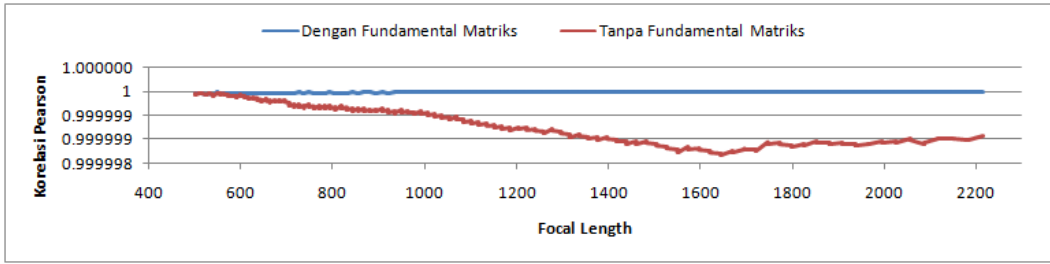


Gambar 4.82: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 1 piksel.

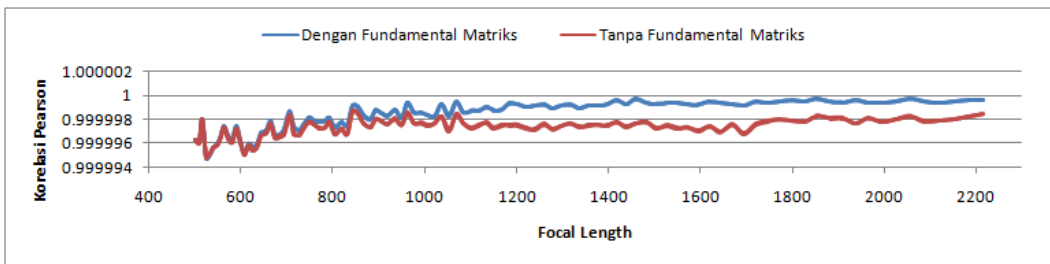
Gambar 4.79 hingga Gambar 4.82 menunjukkan perbandingan MSE hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual 2. Pada 4.79 hingga Gambar 4.81 menunjukkan bahwa hasil rekonstruksi yang melibatkan estimasi menggunakan fundamental matriks memberikan hasil yang lebih mendekati *ground truth* hingga pada akurasi 0.1 piksel. Pada 4.82, hasil rekonstruksi yang tidak melibatkan fundamental matriks memberikan hasil yang lebih baik meskipun Gambar 4.56 dan 4.64 menunjukkan hasil rekonstruksi mengalami distorsi.



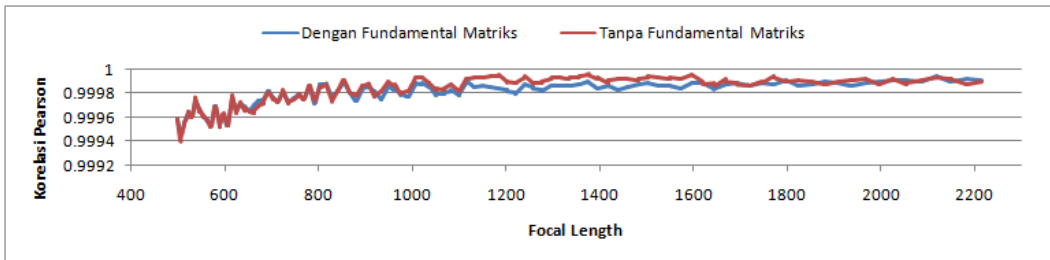
Gambar 4.83: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi penuh.



Gambar 4.84: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.01 piksel.



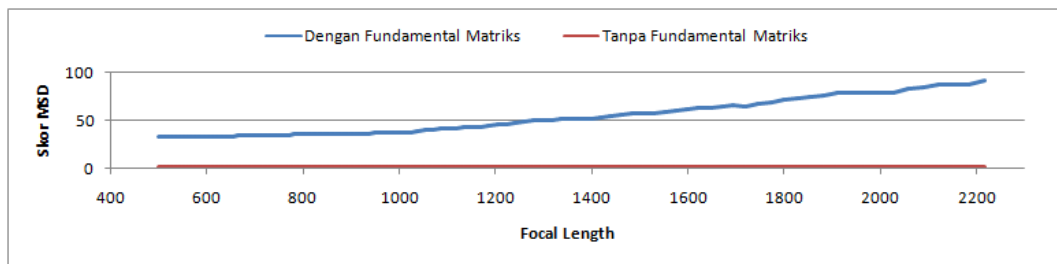
Gambar 4.85: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 0.1 piksel.



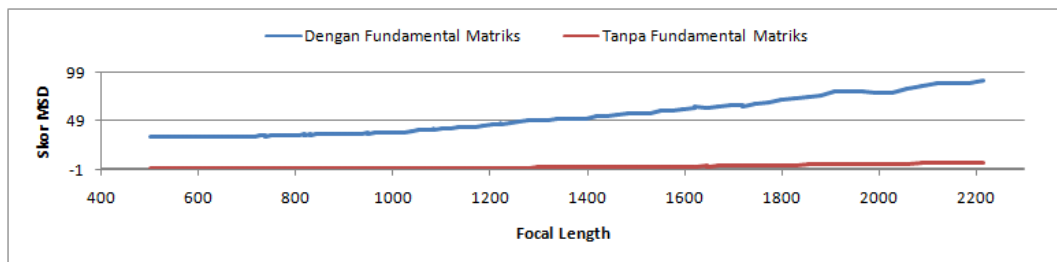
Gambar 4.86: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 2 pada simulasi akurasi 1 piksel.

Gambar 4.83 hingga Gambar 4.86 menunjukkan perbandingan Pearson hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual 2. Pada 4.83 hingga Gambar 4.85 menunjukkan bahwa hasil rekonstruksi yang melibatkan estimasi menggunakan fundamental matriks memberikan hasil yang lebih sebangun dengan *ground truth* hingga pada akurasi 0.1 piksel. Pada Gambar 4.86, kesebangunan hasil rekonstruksi kedua metode relatif setara atau tingkat distorsi bentuk dari hasil rekonstruksi kedua metode setara.

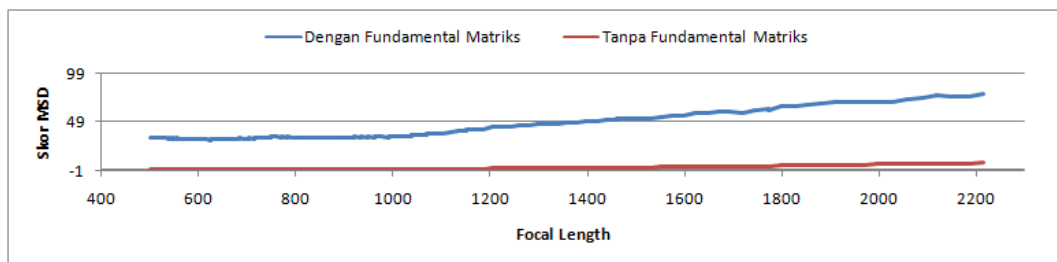




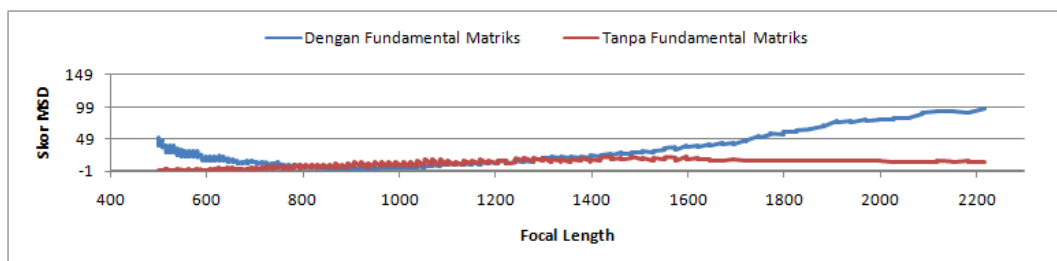
Gambar 4.87: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi penuh.



Gambar 4.88: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.01 piksel.



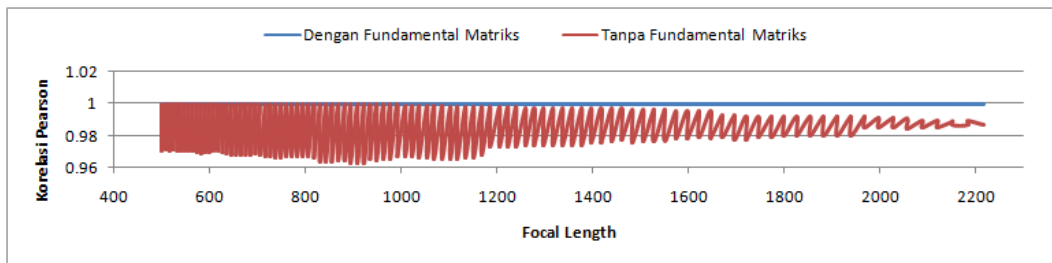
Gambar 4.89: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.1 piksel.



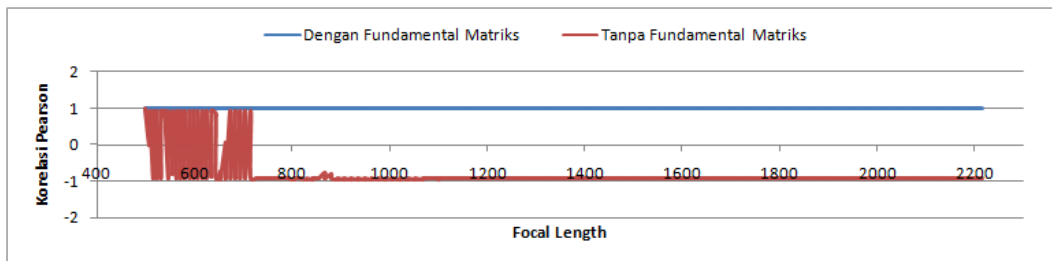
Gambar 4.90: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 1 piksel.

Gambar 4.87 hingga Gambar 4.90 menunjukkan perbandingan MSE hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual

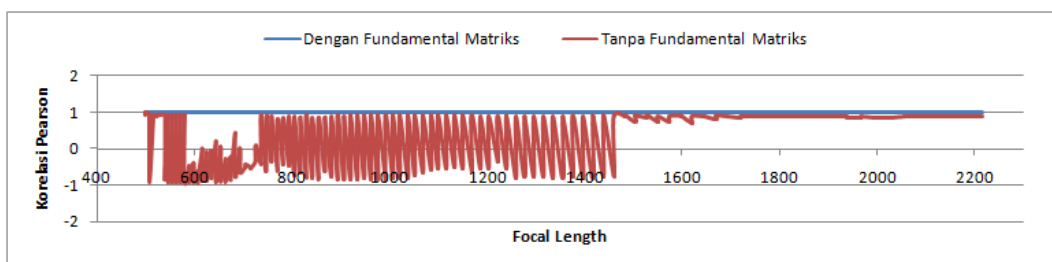
3. Pada 4.87 hingga Gambar 4.89 menunjukkan bahwa skor MSE pada metode estimasi berbasis fundamental matriks selalu lebih tinggi. Hal ini menjadi indikasi hasil rekonstruksi yang tidak menggunakan fundamental matriks memberikan hasil yang lebih mendekati *ground truth* hingga pada akurasi 0.1 piksel. Pada 4.90, hasil rekonstruksi yang tidak melibatkan fundamental matriks memberikan hasil yang lebih baik ketika mencapai *zoom* maksimal namun selama transisi hasil rekonstruksi berbasis fundamental matriks dapat lebih mendekati *ground truth*.



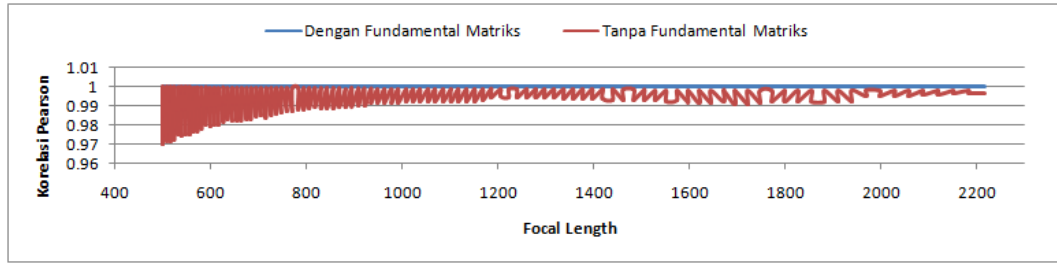
Gambar 4.91: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi penuh.



Gambar 4.92: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.01 piksel.

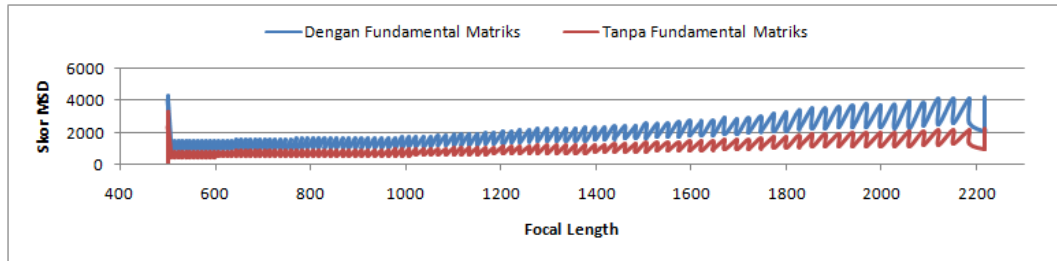


Gambar 4.93: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 0.1 piksel.

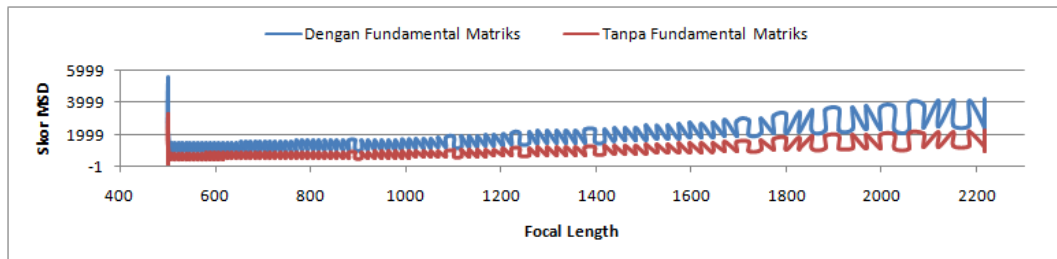


Gambar 4.94: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 pada simulasi akurasi 1 piksel.

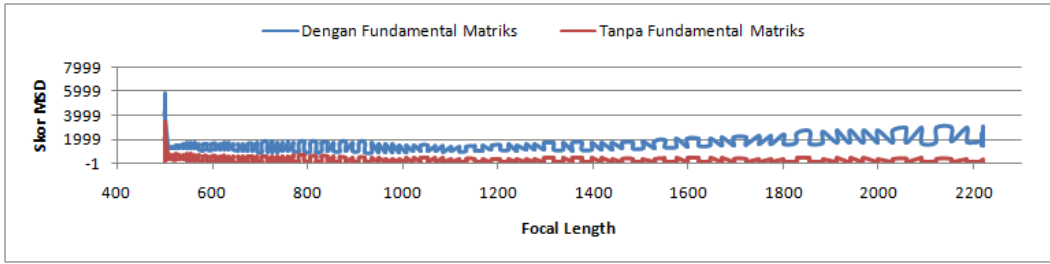
Gambar 4.91 hingga Gambar 4.94 menunjukkan perbandingan Pearson hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual 3. Seluruh skor pearson pada Gambar 4.91 hingga Gambar 4.94 menunjukkan bahwa hasil rekonstruksi yang melibatkan estimasi menggunakan fundamental matriks memberikan hasil yang lebih sebangun dengan *ground truth* meskipun selisih ukuran hasil rekonstruksi dengan *ground truth* *membesar*.



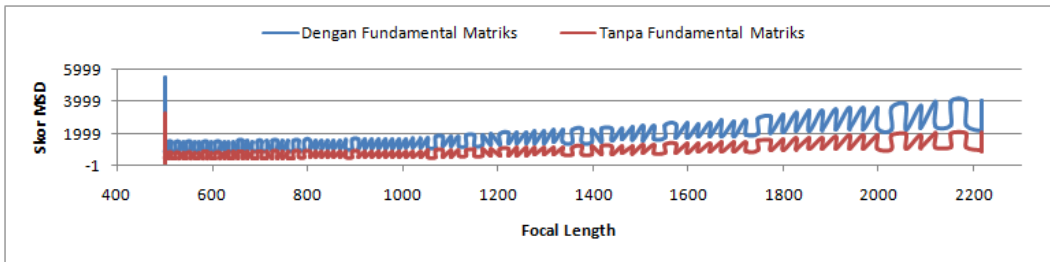
Gambar 4.95: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi penuh.



Gambar 4.96: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi 0.01 piksel.



Gambar 4.98: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi 1 piksel.

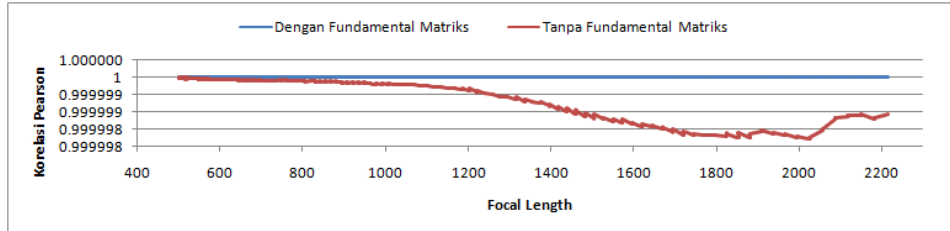


Gambar 4.97: Perbandingan evaluasi MSE hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi 0.1 piksel.

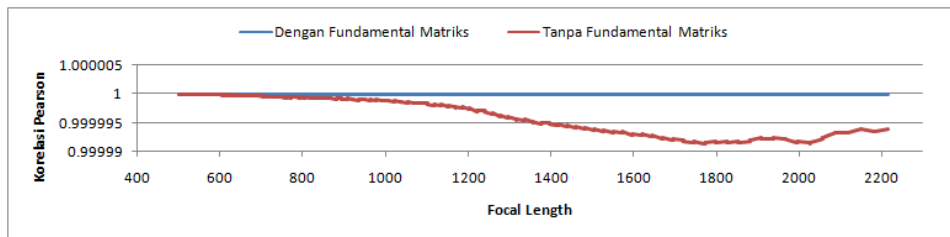
Gambar 4.95 hingga Gambar 4.98 menunjukkan perbandingan MSE hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual 3 dengan *image center* dinamis. Dapat dilihat bahwa skor MSE dari hasil rekonstruksi dari metode yang memanfaatkan fundamental matriks lebih tinggi selalu lebih tinggi pada seluruh level simulasi akurasi. Hal ini menjadi indikasi hasil rekonstruksi yang tidak menggunakan fundamental matriks memberikan hasil yang lebih mendekati *ground truth* meskipun kedua metode memberikan hasil rekonstruksi yang mengalami distorsi.

Gambar 4.99 hingga Gambar 4.102 menunjukkan perbandingan Pearson hasil rekonstruksi menggunakan kedua metode estimasi dari kamera stereo virtual 3 dengan *image center* dinamis. Skor pearson pada Gambar 4.99 hingga Gambar 4.101 menunjukkan bahwa hasil rekonstruksi yang melibatkan estimasi menggunakan fundamental matriks memberikan hasil yang lebih sebangun dengan *ground truth* hingga pada ketelitian 0.1 piksel meskipun memiliki skor MSE yang lebih besar hasil rekonstruksi dengan *ground truth* *membesar*. Pada 4.102 hasil rekonstruksi tanpa menggunakan fundamental ma-

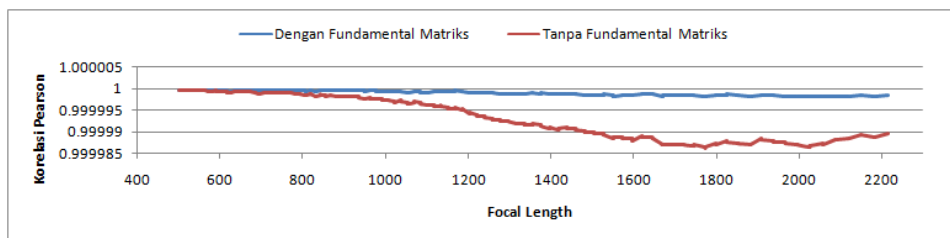
triks lebih sebangun karena skor pearson lebih mendekati 1. Hal ini menjadi indikasi bahwa rekonstruksi menggunakan fundamental matriks dapat mempertahankan bentuk dengan akurasi korespondensi tidak lebih rendah dari 0.1 piksel.



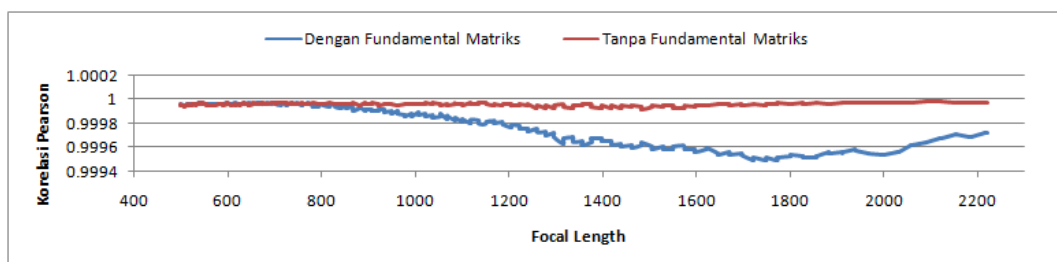
Gambar 4.99: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi penuh.



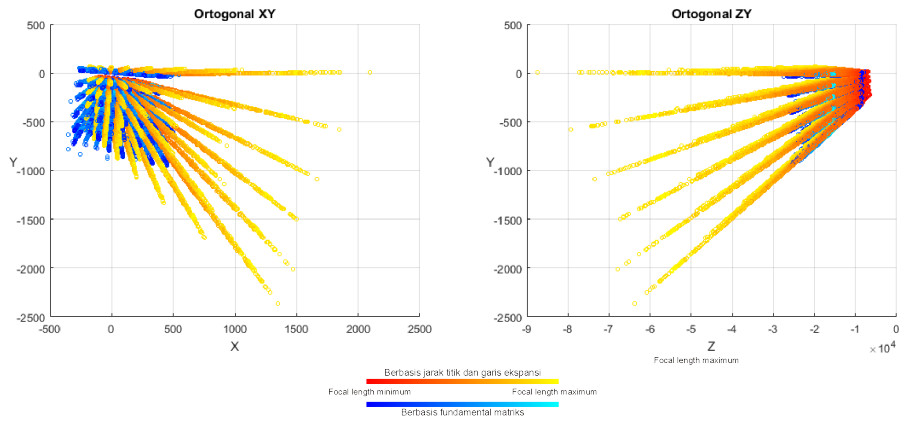
Gambar 4.100: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi 0.01 piksel.



Gambar 4.101: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi 0.1 piksel.



Gambar 4.102: Perbandingan evaluasi Pearson hasil rekonstruksi dari kamera stereo virtual 3 dengan *image center* dinamis pada simulasi akurasi 1 piksel.



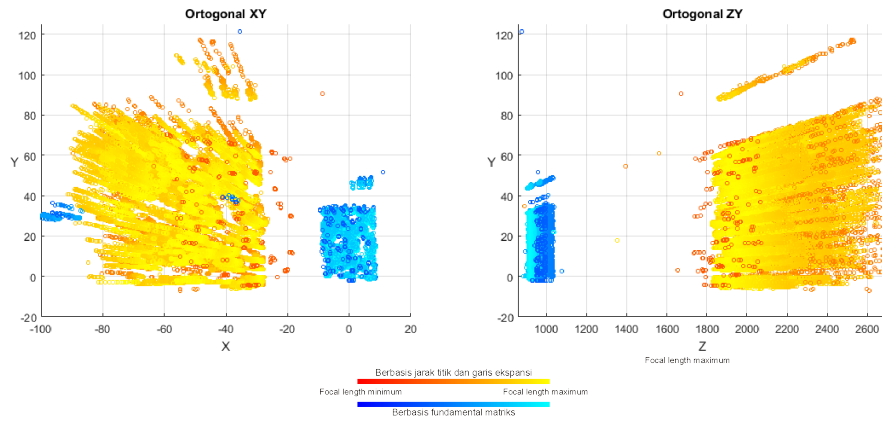
Gambar 4.103: Perbandingan titik-titik hasil rekonstruksi pada scene pertama.

Dari evaluasi data sintetis ini penulis menyimpulkan bahwa hasil rekonstruksi dengan normalisasi berbasis fundamental matriks cenderung lebih akurat ketika akurasi korespondensi mencapai 0.1 piksel. Ketika akurasi korespondensi hanya pada level piksel, hasil rekonstruksi lebih baik jika menggunakan normalisasi berbasis jarak titik dan garis ekspansi. Namun secara keseluruhan hasil rekonstruksi dengan normalisasi berbasis fundamental matriks lebih baik dalam mempertahankan bentuk dari objek.

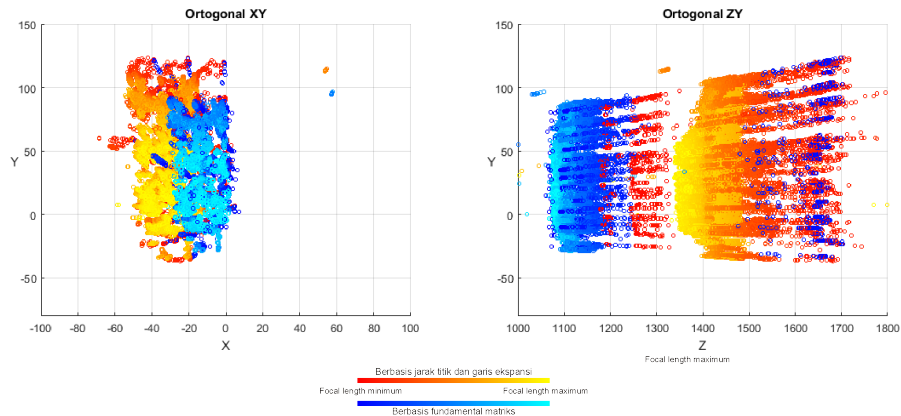
#### 4.6.2. Evaluasi pada data citra

Pada jenis data citra penelitian ini melakukan perbandingan hasil evaluasi rekonstruksi dari kedua jenis metode pada masing-masing *scene*. Perbandingan dilakukan dengan melakukan proyeksi ortogonal dari hasil rekonstruksi pada bidang sumbu  $xy$  dan  $xz$ . Hasil perbandingan disajikan pada Gambar 4.103 hingga Gambar 4.106.

Pada Gambar 4.103 dapat dilihat bahwa persebaran titik-titik hasil rekonstruksi dengan metode normalisasi berbasis fundamental matriks lebih kecil dibandingkan dengan metode normalisasi berbasis jarak titik dan garis ekspansi. Persebaran yang tampak adalah searah sumbu  $x$ ,  $y$  dan  $z$  namun mayoritas persebaran adalah searah sumbu  $z$ . Karena titik-titik berkorespondensi yang digunakan selalu mengacu pada titik tiga dimensi yang sama pada setiap frame, dapat disimpulkan bahwa rekonstruksi dengan metode normalisasi



Gambar 4.104: Perbandingan titik-titik hasil rekonstruksi pada scene kedua.

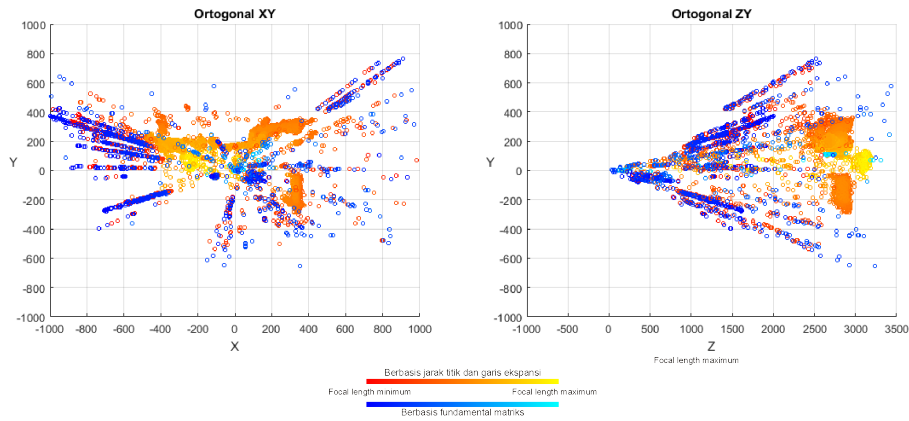


Gambar 4.105: Perbandingan titik-titik hasil rekonstruksi pada scene ketiga.

berbasis fundamental matriks memberikan hasil yang lebih konstan.

Pada Gambar 4.104 dapat dilihat bahwa persebaran titik-titik hasil rekonstruksi dengan metode normalisasi berbasis fundamental matriks lebih kecil dibandingkan dengan metode normalisasi berbasis jarak titik dan garis ekspansi. Persebaran yang tampak adalah searah sumbu  $x$ ,  $y$  dan  $z$  namun mayoritas persebaran adalah searah sumbu  $z$ . Secara metris hasil rekonstruksi berbasis jarak titik dan garis ekspansi memberikaan hasil yang lebih baik namun bila mengamati bentuk dari hasil rekonstruksi objek, rekonstruksi berbasis fundamental matriks memberikan bentuk yang lebih sesuai.

Pada Gambar 4.105 dapat dilihat bahwa persebaran titik-titik hasil rekonstruksi dengan metode normalisasi berbasis fundamental matriks lebih kecil dibandingkan dengan metode normalisasi berbasis jarak titik dan garis ekspansi. Persebaran yang tampak adalah searah sumbu  $x$ ,  $y$  dan  $z$  namun may-



Gambar 4.106: Perbandingan titik-titik hasil rekonstruksi pada scene keempat.

oritas persebaran adalah searah sumbu  $z$ . Secara metris hasil rekonstruksi berbasis jarak titik dan garis ekspansi memberikan hasil yang lebih baik namun bila mengamati bentuk dari rasio antara lebar dan tinggi botol, rekonstruksi berbasis fundamental matriks memberikan bentuk yang lebih sesuai.

Pada Gambar 4.106, hasil rekonstruksi berbasis fundamental matriks tidak menunjukkan bentuk dari objek apapun di dalam *scene*. Hasil rekonstruksi berbasis jarak titik dan garis ekspansi berhasil menunjukkan beberapa objek di *scene* luar ruangan

Dari evaluasi data sintetis ini dapat disimpulkan bahwa hasil rekonstruksi dengan normalisasi berbasis fundamental matriks cenderung lebih mampu mempertahankan bentuk objek dan lebih konstan pada posisi tertentu dilihat dari kecilnya persebaran titik hasil rekonstruksi. Pada rekonstruksi objek berbasis jarak titik dan garis ekspansi, posisi dan ukuran cenderung lebih akurat.



## BAB 5

### KESIMPULAN

Pada penelitian ini, metode estimasi *focal length* dan *center of expansion* pada kamera stereo berbasis fundamental matriks diusulkan. Metode ini sanggup mendapatkan nilai estimasi faktor skala antar frame yang ditangkap oleh kamera stereo yang mengalami proses *zooming*. Faktor skala antara citra berkorelasi langsung rasio *focal length*. Dengan meningkatnya *focal length*, akurasi bentuk hasil rekonstruksi meningkat. Selain itu metode ini juga mampu memberikan estimasi *center of expansion* untuk keperluan normalisasi pada bidang citra terkalibrasi.

Penelitian ini membandingkan metode ini dengan metode estimasi berbasis jarak antar titik dan garis ekspansi. Metode estimasi berbasis jarak antar titik merupakan dasar dari penentuan faktor skala. Metode berbasis jarak antar titik memerlukan pencarian korespondensi antar frame dari kamera yang mengalami proses *zooming*. Dengan demikian, pada kasus kamera stereo pencarian korespondensi pada masing-masing kamera dilakukan secara terpisah. Metode estimasi *center of expansion* berbasis garis ekspansi digunakan atas dasar semua titik yang mengalami skala bergerak pada garis yang menghubungkan titik-titik dengan *center of expansion*. Metode ini memiliki beberapa kemungkinan jawaban dengan posisi vertikal yang berbeda.

Rata-rata hasil estimasi berbasis fundamental matriks pada kondisi *scene* yang dikondisikan mengandung pola papan catur 3.42% lebih besar dibandingkan dengan estimasi berbasis jarak antar titik dengan koefisien *pearson correlation* 0.999. Pada *scene* luar ruangan yang tidak mengandung objek khusus, rata-rata hasil estimasi berbasis fundamental matriks 8.72% lebih kecil dari hasil estimasi berbasis jarak antar titik. Hasil dari penggunaan data sintetis menunjukkan bahwa hasil estimasi *focal length* cukup akurat pada berbagai akurasi korespondensi titik.

Pada jenis data yang disimulasikan metode estimasi berbasis fundamental matriks memberikan hasil rekonstruksi yang lebih sesuai dengan proporsi

dan bentuk objek dengan rata-rata koefisien pearson 0.999999726 sedangkan metode berbasis jarak titik dan garis ekspansi memiliki koefisien 0.999997134. Namun, metode estimasi berbasis jarak titik mampu memberikan ukuran dan posisi yang lebih akurat dengan rata-rata skor MSE sebesar 812.3419562 sedangkan metode berbasis fundamental matriks memiliki skor 1774.45438. Pada jenis data citra fundamental matriks memberikan hasil rekonstruksi yang memiliki simpangan yang lebih pendek dibandingkan metode berbasis jarak antar titik dan garis ekspansi.

Metode yang diusulkan pada penelitian ini mampu melakukan rekonstruksi tiga dimensi dari sistem kamera stereo dengan *focal length* dinamis. Penelitian ke berikutnya dapat dilakukan dengan menggunakan dua kamera dengan transformasi antar kamera mengandung rotasi. Dengan adanya rotasi, garis epipolar pada citra belum tentu horizontal dan pergeseran ketika *zooming* bisa berbeda.

## DAFTAR PUSTAKA

- Bhatti, A. & Navahandi, S. (2011), *ADVANCES IN THEORY AND APPLICATIONS OF STEREO VISION*, InTech, Rijka, Croatia.
- Elamsy, T., Habed, A. & Boufama, B. (2014), ‘Self-calibration of stationary non-rotating zooming cameras’, *Image and Vision Computing* **32**(3), 212–226.  
**URL:** <http://dx.doi.org/10.1016/j.imavis.2014.01.003>
- Hartley, R. & Zisserman, A. (2002), ‘Multiple view geometry in computer vision’, **37**, 85–86.
- Khilar, R., Chitrakala, S. & SelvamParvathy, S. (2013), ‘3D image reconstruction: Techniques, applications and challenges’, *2013 International Conference on Optical Imaging Sensor and Security (ICOSS)* pp. 1–6.  
**URL:** <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6678395>
- Li, M. & Lavest, J.-M. (1996), ‘Some Aspects of Zoom-Lens Camera Calibration’, *IEEE Trans on Pattern Analysis and Machine Intelligence* **18**(11), 1105–1110.
- M. Pollefeys, R. K. & Gool, L. V. (1998), ‘Self-Calibration a8nd Metric Reconstruction in spite of Varying and Urilmown Internd Chrriera Pa,r.anieters’, *Sixth International Conference on Computer Vision* pp. 90–95.
- Nalpantidis, L., Amanatiadis, A., Sirakoulis, G., Kyriakoulis, N. & Gasteratos, A. (2009), ‘Dense disparity estimation using a hierarchical matching technique from uncalibrated stereo vision’, *2009 IEEE International Workshop on Imaging Systems and Techniques* pp. 427–431.  
**URL:** <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5071680>
- Vancea, C. C., Miclea, V. C. & Nedevschi, S. (2016), ‘Improving stereo reconstruction by sub-pixel correction using histogram matching’, *IEEE Intelligent Vehicles Symposium, Proceedings 2016-Augus(Iv)*, 335–341.

- Willson, R. G. (1994), ‘Modeling and Calibration of Automated Zoom Lenses’, *Proceedings of the {\sc spie} Conference on Videometrics III, Boston, Massachusetts, USA* **2350**(January), 170–186.
- Zollner, H. & Sablatnig, R. (2004), ‘Comparison of methods for geometric camera calibration using planar calibration targets’, *Proceedings of the 28th Workshop of the Austrian Association for Pattern Recognition* pp. 237–244.

## BIOGRAFI PENULIS



Penulis bernama Kurnia Prima Putra lahir di Malang pada 18 November 1990. Penulis merupakan putra sulung dari Ir. Pudji Mulyo Winarno (Alm.) dan Dra. Ety Tejo Dwi Cahyowati, M.Pd. Pada masa kecilnya, penulis menimba ilmu di SD Sriwedari Malang dimulai tahun 1996 dan lulus pada tahun 2002. Selanjutnya penulis mengenyam pendidikan menengah pertama di SMP Negeri 1 Malang pada tahun 2002 hingga 2005 dan dilanjutkan ke pendidikan menengah atas pada tahun 2005 hingga 2008. Pada tahun 2008 penulis melanjutkan pada pendidikan tinggi tingkat Sarjana di Universitas Brawijaya pada bidang keahlian Teknik Informatika dan berhasil memperoleh gelar Sarjana pada tahun 2014 dengan topik skripsi arsitektur perangkat lunak. Pada tahun 2016 penulis melanjutkan pendidikan pada jenjang Magister di Institut Teknologi Surabaya.

Penulis dapat dihubungi melalui email [kprimap@gmail.com](mailto:kprimap@gmail.com).

